FACULTEIT ECONOMIE EN BEDRIJFSWETENSCHAPPEN

KU LEUVEN

THE WORLD BEYOND GARP: METHODOLOGICAL ADVANCES IN REVEALED PREFERENCE THEORY

Proefschrift voorgedragen
tot het behalen van de graad
van Doctor in de Economische
Wetenschappen

door

Sam COSAERT

Serial Number 483

2015

Daar de proefschriften in de reeks van de Faculteit Economische en Toegepaste Economische Wetenschappen het persoonlijk werk zijn van hun auteurs, zijn alleen deze laatsten daarvoor verantwoordelijk.

Doctoral committee

Supervisor:	Prof. dr. Laurens Cherchye
Co-supervisor:	Prof. dr. Thomas Demuynck
Members:	Prof. dr. Ian Crawford
	Prof. dr. Bram De Rock
	Prof. dr. Mark Dean
	Prof. dr. Geert Dhaene
	Prof. dr. Frederic Vermeulen
Chairman:	Prof. dr. Frans Spinnewyn

Acknowledgements

You never know what's around the corner. It could be everything. Or it could be nothing. You keep putting one foot in front of the other, and then one day you look back and you've climbed a mountain.

Tom Hiddleston

When I started my PhD on revealed preferences almost 4 years ago, I was unaware of how the research project would unfold and where it would lead me. However, I was fortunate to have excellent guides to help me proceed, step by step. Now that I am looking back at the PhD process, it is appropriate to show my gratitude to these people who helped me 'climb the mountain'.

First and foremost, I would like to thank my supervisor Prof. Laurens Cherchye. I have been lucky to have a supervisor who cared so much about my work. During the four years of my PhD, Laurens' gentle encouragement and enthusiasm for new ideas have been truly unique. His ability to look beyond obstacles typically revealed more interesting research questions or more promising approaches. Laurens shared not only his intellectual expertise in research, he also encouraged me to explore other aspects of the Life in Academia, including workshops and conferences. I thank Laurens for these opportunities.

This PhD would not have been possible without the support of my co-supervisor Prof. Thomas Demuynck. Thomas combines his extraordinary talent with uncommon modesty.

I have been greatly impressed by Thomas' genius and his ability to develop new ideas. By cooperating with him, I was able to develop my scientific skills and improve my understanding of the revealed preference topic. I also greatly appreciated Thomas' outright and sincere advice.

A special thanks goes to Prof. Bram De Rock. Together with Laurens, he introduced the topic of revealed preference to me during the second year of my bachelor education. He was also one of the first people with whom I discussed the possibility of pursuing a PhD. I thank Bram for conveying his enthusiasm for the revealed preference topic and for the time and effort invested in our joint work.

I am truly honoured that Prof. Frederic Vermeulen, Prof. Geert Dhane, Prof. Ian Crawford and Prof. Mark Dean were willing to accept our invitation to become members of my doctoral committee. Their thoughtful suggestions significantly improved the quality of this dissertation. I also greatly appreciate the opportunity offered by Prof. Crawford to stay at Nuffield College in Oxford. This stay has widened my perspective on nonparametric demand analysis and has made me think more critically of my own research.

Furthermore, I am grateful to Trui Stragier for the management of my conference visits and to Annie Vercruysse for the management of my PhD progress.

I thank my colleagues Veerle and Pieter Jan for sharing the happy and not so happy moments associated with the PhD process during 3 and 2 years, respectively. When the temperature in our office C712 got too high, I enjoyed the refreshing walks along the 'Spina' to get something to drink. I also thank Ewout, who guided me through MATLAB and through the first year of my PhD, and Jeroen, whose contributions to the research team were very valuable to me. Furthermore, I thank Charlotte and Griet for their interesting fashion and life style advice (and for their assistance in the quest for Joeri's new jacket) and Anneline, Lindsay, Johan, Marijn, Dries, Jan-Bart and Stephan for their appreciated company during lunch time and coffee breaks. I also thank my former colleagues Simon, Thijs, Sofie, Gaëlle, Skrålan and Brecht. Finally, I wish to express my respect to André, who has been very supportive during my PhD and the preceding bachelor years. He has always believed in me and in the successful completion of this PhD project. I thank my colleagues from Leuven, and in particular, the members of the public economics research group. I thank Wouter for the enjoyable chats at the 'Stuk'. I also thank the participants to the reading group on collective consumption models, and in particular, Denni, Alex and Tom.

I wish to express my love and gratitude to Joeri. Our humorous conversations - one of the reasons that we became friends twelve years ago - have been ideal to cope with occasional stressful moments. It is truly remarkable to meet someone so open-hearted and accepting. Joeri, I could always rely on you, especially when I needed encouragement. Although it is uncertain where the road will lead us from here, I am not worried. I enjoy being together with you, whether it be in Kortrijk, Leuven, Tiegem or somewhere else. I also thank our family members and friends for their continuing love. Our thoughts are with oma Christine, who would have been very proud to hold this dissertation. Her cheerful enjoyment of life will be missed.

Most of all, I thank my parents. They provided me with a warm-hearted and supportive environment. They taught me to be perseverant, patient and humble, yet confident. I really enjoy the good talks during our weekly visits, especially when these visits take place in the 'Kudu Lodge', the 'Irish Pipes' or 'Sachsen'. I am looking forward to the adventures yet to come, including our hike along the notorious Malerweg. Although these words of thanks cannot fully express my gratitude, it feels good to finally show my appreciation for their support.

At last, I would like to thank you, the reader, for showing interest in this dissertation. I have genuinely enjoyed working on the topic of consumer choice and I sincerely hope that I can convey some of my enthusiasm in this dissertation.

Contents

List of Tables

List of Figures

Part I

General Introduction

In January 2005, Varian, one of the most influential contributors to the revealed preference literature, ran a search for 'revealed preference' on Jstor and Google Scholar, and found 997 and 3600 papers, respectively. In August 2014, I repeated this exercise and found up to 1035 new results on Jstor and 17400 new results on Google Scholar since 2005. This clearly indicates that the revealed preference method is a highly relevant and popular approach.

In this doctoral dissertation, I formulate methodological extensions of the revealed preference approach founded by Samuelson (1938, 1948) and Houthakker (1950). The revealed preference approach allows us to impose consistency on observed choices from (usually linear) budget sets[1]. First of all, I will modify the standard revealed preference principles to *test consistency* of choices from finite choice sets. Second, I will generalise the narrowly defined revealed preference axioms (GARP in particular) to incorporate psychological realism in the models. Specifically, I will *identify* preferences for others' consumption and preferences for value (diamond effects). Finally, I will show how revealed preference techniques can help us estimate the distribution of welfare measures and *predict* the distribution of demand correspondences in a setting where panel data is unavailable.

In my General Introduction, I will consider the foundations of revealed preference theory by Samuelson (1938, 1948) and Houthakker (1950). Then I will discuss the potential of revealed preference theory from a methodological perspective. Seminal contributions by Afriat (1967), Diewert (1973) and Varian (1982) will be briefly explained. Moreover, I will argue that the revealed preference method can not be confined to one theory or one application. The method can be used to study a large variety of research questions. Finally, I will discuss the main themes in the revealed preference literature (testing, identification and prediction) and position my results in this framework.

[1]Note that in particular settings (e.g. experimental and transportation related studies) the term 'revealed preference' has a slightly different connotation. It may refer to the observation of real market decisions, as opposed to letting respondents choose in hypothetical situations, without (implicitly or explicitly) imposing rationality.

Background Paul Samuelson established the foundations of revealed preference in his 1938 Economica article. In order to understand the intentions of Paul Samuelson, it is important to understand the evolution of the economic science until 1938.

In the nineteenth century, there was serious debate among economists whether the economic science should move towards political economy, putting emphasis on values and ethics, or towards a more empirical (positivist) approach, putting the economic science on par with natural sciences. John Stuart Mill, as well as John Neville Keynes, argued in favour of the distinction between positive and normative economics. Keita (2012) pointed out that these economists were most likely influenced by the empiricist views developed by Locke and Hume.

However, studying economic decisions in a similar way as chemistry and physics study natural phenomena appeared to be rather ambitious, if not problematic. After all, social sciences like economics study real humans, whose behaviour is intrinsically different from natural phenomena. One crucial difference is that human actions are based on internal considerations about their own well-being. To successfully create theories of human behaviour, researchers needed a measure for well-being. This led to the *utilitarian* approach, founded by Jeremy Bentham (1879) and James Mill. Utilitarianism is based on the principle that people's well-being can be measured in terms of individual pleasure and pain (the *felicific calculus of pleasure and pain*).

By the end of the nineteenth century, economists like Cournot, Dupuit and Gossen started to develop theories based on individual utilities. The true neo-classical revolution began around 1870. Jevons (1879) published *Theory of Political Economy*, in which he observed that the marginal utility (the *final degree of utility*) associated with some good is decreasing in the level of consumption of this good. More generally, the neo-classical approach deviated from the classical approach in that 1) the purpose of all observed behaviour is assumed to be the maximisation of individual utilities, 2) more attention is given to changes in utilities (the *marginalist paradigm*) and 3) decision-makers are atomic elements such as individuals, households or firms, rather than groups of agents from a certain economic or social class. Despite its potential to formulate economic theories in rigorous mathematical terms, the utilitarian approach experienced failure. First of all, it was argued that individual utilities are not measurable, as they are the result of internal considerations on the account of the consumer. Second, even if there was an objective measure for utility, it is practically impossible to observe levels of utility.

For these reasons, Hicks and Allen (1934), Pareto (1927) and Slutsky (1915) started to treat utility as an ordinal concept rather than a cardinal measure. This 'escape from psychology' (Giocoli (2003)) relaxed the assumptions that utility is cardinally measurable, and that it can be summed across various individuals. Instead, these authors came up with ordinal utility theory, which assumes that each individual can rank alternatives according to his or her preferences.

In 1938, Paul Samuelson attempted to go one step further. He argued that, on the one hand, many assumptions on the preferences of individuals had been dropped since the time of Gossen. Two notable examples of such assumptions are the linearity of marginal utility and the measurability of utility in a cardinal sense. After all, ordinal utility theory restricts itself to the analysis of indifference elements and the relationship between relative prices and the slopes of indifference curves. On the other hand, Samuelson (1938) observed that the ordinal utility theory still relied on assumptions which were hard to verify. The theory typically assumes that the marginal rate of substitution is increasing. Samuelson (1938) proposed to replace the ordinal utility theory with a more direct approach. Instead of putting restrictions on preferences, which are unobserved, Samuelson put restrictions on demand, which is observable. This contribution can be seen as the foundation of the revealed preference approach.

Foundations In his 1938 Economica article, Samuelson considered the choices made by an 'idealised' homo economicus. The revealed preference logic, as explained by Samuelson

(1938), is as follows:

> 'If this cost [the cost associated with prices of one position applied to the batch of goods bought in a second position] is less than or equal to the actual expenditure in the first period when the first batch of goods was actually bought, then it means that the individual could have purchased the second batch of goods with the price and income of the first situation, but did not choose to do so. That is, the first batch (x) was selected over [the second batch] x".

This clearly illustrates how revealed preference relations (the first bundle is preferred to the second bundle) are constructed. Quite remarkably, the construction of revealed preference relations has hardly changed over the past decades. Samuelson (1938) proceeded by formulating axioms on the demands (rather than the preferences) of consumers. Specifically, he postulated that the 'idealised' individual's behaviour is consistent. Consistency means that if the first batch was selected over the second, then the second batch can not be selected simultaneously over the first. Given price vectors $\mathbf{p}_t$ and quantity vectors $\mathbf{q}_t$ associated with different observations $t, v \in T$, let us formalise the (weak) direct revealed preference relation R_0 :

$$\mathbf{p}_t'\mathbf{q}_t \geq \mathbf{p}_t'\mathbf{q}_v \Rightarrow \mathbf{q}_t \; R_0 \; \mathbf{q}_v$$

Then Samuelson's *Weak Axiom of Revealed Preference* indicates that if $\mathbf{q}_t \; R_0 \; \mathbf{q}_v$, we can not simultaneously have that $\mathbf{q}_v \; R_0 \; \mathbf{q}_t$. At the time, Samuelson (1938) presented his revealed preference approach as an alternative to ordinal utility theory, hoping to get rid of the utility concept all together. However, by 1948, Samuelson appears to recognise that his revealed preference theory and ordinal utility theory are complementary. Following Little (1949) he used revealed preference to construct 'revealed preferred' or 'revealed worse' regions in an indifference map and to approximate the indifference curve. It appears that Samuelson's research program had evolved from replacing ordinal utility theory towards using revealed preference to empirically verify concepts of the ordinal utility theory (Samuelson (1948)).

One of the unanswered questions in Samuelson's 1938 Economica article was whether the Weak Axiom of Revealed Preference (WARP henceforth) exhibits all of the implications associated with the utility maximisation hypothesis. It was clear that the postulate of utility maximisation implied the WARP, while it was uncertain whether WARP implied consistency with a well-behaved preference ordering. In 1950, Houthakker showed that WARP does not capture all aspects of the utility maximisation hypothesis. Houthakker (1950) strengthened the WARP by introducing transitivity of demand and indirect revealed preference relations R. For each sequence of bundles $\mathbf{q}_a, \mathbf{q}_b, ..., \mathbf{q}_s$ we have that

$$\mathbf{q}_t \; R_0 \; \mathbf{q}_a \; , \mathbf{q}_a \; R_0 \; \mathbf{q}_b \; , \ldots , \mathbf{q}_s \; R_0 \; \mathbf{q}_v \;\Rightarrow\; \mathbf{q}_t \; R \; \mathbf{q}_v$$

Houthakker's *Strong Axiom of Revealed Preference* indicates that if $\mathbf{q}_t \; R \; \mathbf{q}_v$, we can not simultaneously have that $\mathbf{q}_v \; R_0 \; \mathbf{q}_t$. The Strong Axiom of Revealed Preference (SARP henceforth) exhibits all of the testable implications associated with the utility maximisation model.

Extended and restricted domain versions of revealed preference Summarising, the revealed preference approach evolved from being 'a possible alternative to ordinal utility theory', over 'complementary to ordinal utility theory' to 'equivalent to ordinal utility theory'. Many researchers have described this as the failure of the revealed preference research program. Fortunately, this did not stop the revealed preference approach from moving forward. Let me discuss two promising ways in which revealed preference research proceeded after 1950.

First of all, Arrow (1959), Richter (1966), Sen (1971) and Suzumura (1976) made progress by extending and generalising the existing methods. Pollak (1990) referred to this research as the *extended domain* version of revealed preference. Arrow (1959) generalised Samuelson's work in two ways. On the one hand, he incorporated the possibility that decision makers choose more than one alternative. On the other hand, the author allowed for more general sets of alternatives. Indeed, the theory had long been confined to analysing demand from budget sets (budget triangles in a two-goods case). In his paper, Arrow (1959) allowed for choices made from traditional budget sets but also from two-element sets and finite choice sets. The corresponding preference orderings build on binary comparisons of alternatives. This generalisation was far from trivial, because it paved the way for applications of revealed preference theory to decisions by, for instance, governments (see for instance Basu (1980) on revealed preference of government). Rather than choosing between various bundles from budget triangles, governments typically choose from a finite set of alternative projects. In this context, authors have come up with alternative axioms on observed behaviour (such as *Weak* and *Strong Congruence* by Richter (1966)) and alternative rationality concepts (such as *G-rationality* by Suzumura (1976) or *regular rationality* by Richter (1966)). In Chapter 2, I will build the bridge between this extended version of revealed preference theory and standard revealed preference theory in the tradition of Samuelson (1938) and Houthakker (1950). In particular, I consider revealed preference tests for choices from finite choice sets (rather than linear budget sets) while still using standard utility functions.

Second, Afriat (1967), Diewert (1973) and Varian (1982) developed a version of revealed preference for settings with limited information on prices and quantities, i.e. when the set of observed prices and quantities is finite. Pollak (1990) referred to this research as the *restricted domain* version of revealed preference. This is in sharp contrast with the initial contributions of Samuelson and Houthakker. Samuelson and Houthakker, on the one hand, applied their revealed preference principles to the demands of consumers, assuming the observability of a demand system. Afriat (1967), on the other hand, used only a finite set of prices and quantities as inputs. It is worth noting that the two different approaches correspond to two different purposes. While standard domain revealed preference is particularly concerned with 'exhausting the implications of revealed preference theory', the restricted domain version aims at 'testing the utility maximisation hypothesis on the basis of finite sets of prices and quantities' (see Pollak (1990)). Otherwise stated, the restricted domain version is more well-suited as a practical test of consumer behaviour. It allows to verify statements on the rationality of consumers, or recover elements of the underlying preference structure, without imposing structure on utility functions and without assuming the observability of a full demand system. While the equivalence between SARP and the utility maximisation hypothesis seemed to render revealed preference redundant, Afriat (1967), Diewert (1973) and Varian (1982) clearly demonstrated the method's methodological potential.

The methodological contributions of Afriat, Diewert and Varian Afriat (1967) first applied the revealed preference method to a data set $S = \{(\mathbf{p}_t, \mathbf{q}_t) | \forall t \in T\}$ which contains a finite number (T) of price vectors $\mathbf{p}_t$ and quantity vectors $\mathbf{q}_t$. He presented an approach to decide whether an observed data set of type S is *utility consistent*, i.e. whether there exists a utility function that can generate the observed data set. Afriat (1967) concluded that such utility function exists provided that the data set S satisfies the property of *cyclical consistency*. Cyclical consistency is defined as follows:

$$\begin{aligned} & \mathbf{q}_t \; R_0 \; \mathbf{q}_s \; , \mathbf{q}_s \; R_0 \; \mathbf{q}_r \; , \ldots , \mathbf{q}_v \; R_0 \; \mathbf{q}_t \\ \Rightarrow \quad & \mathbf{p}_t' \mathbf{q}_t = \mathbf{p}_s' \mathbf{q}_s = \mathbf{p}_r' \mathbf{q}_r = \mathbf{p}_v' \mathbf{q}_v \end{aligned}$$

Otherwise stated, if some series of direct (and weak) revealed preference relations form a cycle, it must be the case that the expenditures in the corresponding periods are equal. Afriat (1967) proceeded by showing that the cyclical consistency requirement is equivalent to *level consistency,* which is formalised as consistency with a system of linear inequalities (the so called Afriat inequalities).

$$u_t - u_v \leq \lambda_v \mathbf{p}_v' (\mathbf{q}_t - \mathbf{q}_v)$$

If there exist utility *levels* u_t and u_v and strictly positive *multipliers* λ_v such that the inequalities hold for all observations v and t, the data are said to be level consistent. Afriat (1967) used this system of inequalities to re-construct a utility function that could have generated the observed data. Specifically,

$$u(\mathbf{q}) = \min_v u_v + \lambda_v \mathbf{p}'_v(\mathbf{q} - \mathbf{q}_v)$$

is a utility function that 'realises the utility hypothesis on S'. In this way, Afriat (1967)'s method allows us to test consistency with the utility hypothesis (by verifying consistency with a system of linear inequalities) and construct a utility function that imposes consistency on the observed data.

Diewert (1973) clarified Afriat (1967)'s exposition by making assumptions on the utility functions $u(\mathbf{q})$. The author found, for instance, that Afriat's notion of *cyclical consistency* is equivalent to the existence of a locally non-satiated utility function. Moreover, Diewert (1973) found that information on a finite number of choices (from linear budget sets) does not enable researchers to make a distinction between consistency with a locally non-satiated utility function on the one hand, and a locally non-satiated utility function which is also increasing, continuous and concave on the other hand.

Let me finally discuss Varian's contributions to the revealed preference approach. First of all, Varian (1982) developed a characterisation which is formally equivalent to Afriat's cyclical consistency, but which is easier to test in practice. Towards this end, Varian (1982) made use of the strict direct revealed preference relation P :

$$\mathbf{p}'_t\mathbf{q}_t > \mathbf{p}'_t\mathbf{q}_v \Rightarrow \mathbf{q}_t \; P \; \mathbf{q}_v$$

He reformulated cyclical consistency as follows: if $\mathbf{q}_t \; R \; \mathbf{q}_v$, we can not simultaneously have that $\mathbf{q}_v \; P \; \mathbf{q}_t$. In other words, if $\mathbf{q}_t$ is (indirectly) revealed preferred over $\mathbf{q}_v$, it should not be the case that $\mathbf{q}_v$ is strictly directly revealed preferred over $\mathbf{q}_t$. This became known as the *Generalised Axiom of Revealed Preference* (GARP henceforth). It is easily seen that this characterisation is similar in spirit to Samuelson's WARP and Houthakker's SARP. The only difference is that GARP allows for indifference curves with flat parts and SARP does not. Interestingly, Varian (1982) showed that checking consistency with GARP is computationally more efficient than checking consistency with Afriat's inequalities.

Definition 0.1. *GARP*

The set $S = \{(\mathbf{p}_t; \mathbf{q}_t); t = 1, ..., T\}$ is consistent with *GARP* if there exist direct revealed preference relations R_0 and indirect revealed preference relations R such that

1. if $\mathbf{p}_t'\mathbf{q}_t \geq \mathbf{p}_t'\mathbf{q}_v$, then $\mathbf{q}_t R_0 \mathbf{q}_v$;
2. if $\mathbf{q}_t R_0 \mathbf{q}_r, \mathbf{q}_r R_0 \mathbf{q}_s, ...\mathbf{q}_u R_0 \mathbf{q}_v$, then $\mathbf{q}_t R \mathbf{q}_v$;
3. if $\mathbf{q}_t R \mathbf{q}_v$, then $\mathbf{p}_v'\mathbf{q}_t \geq \mathbf{p}_v'\mathbf{q}_v$.

Intuitively, Statement 1 constructs direct revealed preference relations: if $\mathbf{q}_v$ was in the interior of the budget set associated with period t (so that $\mathbf{q}_v$ was affordable at time t although $\mathbf{q}_t$ was chosen), we can say that $\mathbf{q}_t$ is directly revealed preferred over $\mathbf{q}_v$. Statement 2 imposes transitivity on the direct revealed preference relations, thereby constructing indirect revealed preference relations. Finally, Statement 3 imposes expenditure minimisation: if $\mathbf{q}_t$ is equally good or better than $\mathbf{q}_v$, then $\mathbf{q}_t$ should not be strictly cheaper in period v (since otherwise there would be no reason to choose $\mathbf{q}_v$ in period v).

The methodological contributions of Afriat, Diewert and Varian have promoted revealed preference as a useful tool for empirical analysis. In the same way as standard econometrics allowed to test and re-construct (parametric) demand systems, the revealed preference approach allowed to test consistency with the utility hypothesis and re-construct elements of the underlying preference structure without imposing restrictions on demand or utility functions. However, in order to be truly successful, the revealed preference program also needed measures by which the (empirical) performance of the axioms could be assessed. Fortunately, over the past decades, many authors have made progress in this area.

On the one hand, a decent empirical method should not reject some null hypothesis while it is true (**Type I error**). In this respect, Varian (1990) argued that the (exact) revealed preference tests are overly restrictive. It is indeed reasonable to argue that consumers make optimisation errors when maximising their utility. For this reason, Afriat (1967) proposed to replace the requirement of exactly optimising behaviour by the requirement of nearly optimising behaviour. In this way, rationality is not automatically rejected when a small violation of a revealed preference axiom occurs. An alternative definition of GARP is then

Definition 0.2. *e-GARP*

The set $S = \{(\mathbf{p}_t; \mathbf{q}_t); t = 1, ..., T\}$ is consistent with *e-GARP* if there exist direct revealed preference relations $R_0(e)$ and indirect revealed preference relations $R(e)$ such that

1. if $e_t \cdot \mathbf{p}_t'\mathbf{q}_t \geq \mathbf{p}_t'\mathbf{q}_v$, then $\mathbf{q}_t R_0(e) \mathbf{q}_v$;
2. if $\mathbf{q}_t R_0(e)\mathbf{q}_r, \mathbf{q}_r R_0(e)\mathbf{q}_s, ...\mathbf{q}_u R_0(e)\mathbf{q}_v$, then $\mathbf{q}_t R(e)\mathbf{q}_v$;
3. if $\mathbf{q}_t R(e)\mathbf{q}_v$, then $\mathbf{p}_v'\mathbf{q}_t \geq e_v \cdot \mathbf{p}_v'\mathbf{q}_v$.

The values e_t in vector e relax the rationality requirement when $e_t < 1$. Indeed, $e_t < 1$ implies that fewer revealed preference relations $R_0(e)$ and $R(e)$ are constructed and that the expenditure minimisation requirement in Statement 3 becomes weaker. Varian (1990) characterised this efficiency index:

$$e_t = \min_{\mathbf{q}_v R(e) \mathbf{q}_t} \frac{\mathbf{p}_t'\mathbf{q}_v}{\mathbf{p}_t'\mathbf{q}_t}$$

Intuitively, e_t compares the minimum expenditures needed to purchase a bundle $\mathbf{q}_v$ which is revealed preferred over $\mathbf{q}_t$ with the actual cost of observation t. The lower e_t, the more money was wasted in period t. It is often convenient to consider a uniform bound, a fixed e^* that relaxes each budget constraint by the same fraction. I will refer to this goodness-of-fit index as *Afriat's Critical Cost Efficiency Index*.

On the other hand, empirical analysis is concerned with **Type II error**. A test is less sustainable (Type II error is high) when observations from some alternative hypothesis do not lead to the rejection of the null hypothesis. Similarly, if one particular revealed preference test was unable to reject the utility maximisation hypothesis when confronted with random choice patterns (which do not correspond to decisions of real humans), the test would be rather useless. Bronars (1987) formalised a measure for the strength of a revealed preference test, discriminatory power. Discriminatory power measures the extent to which revealed preference tests can detect and reject random choices. There are various methods to simulate random bundles. Bronars' method draws random budget shares from the uniform distribution. The bootstrap approach draws budget shares from the distribution of observed budget shares across the sample and randomly allocates these shares to an individual's consumption in various periods. A 'strong' revealed preference test is then able to reject consistency with the utility maximisation hypothesis when confronted with these random data sets. The power is by now a widely used measure of empirical success in the revealed preference literature.

Critical discussion The revealed preference program is praised because it avoids putting structure on utility functions. As such, the method abstains from imposing unverifiable assumptions on the preferences of consumers. This means that revealed preference allows for a pure test of rationality which is independent of functional form restrictions (thereby avoiding a 'dual' hypothesis). Moreover, it allows to analyse different individuals separately when different prices and consumption choices are observed per individual. As a result, debatable preference homogeneity assumptions are unnecessary. Finally, as argued above, it is straightforward to check consistency with revealed preference conditions. One can simply check whether the conditions are not mutually exclusive, or in more complicated cases, one can formulate a standard linear programming problem. However, the method has also received considerable criticism.

1. First of all, revealed preference tests often lack discriminatory power. It is easy to construct a setting in which no violations of the standard revealed preference axioms (WARP, SARP or GARP) are possible. This occurs when the prices and incomes, associated with various observations, are such that the budget lines completely dominate each other. Otherwise stated, revealed preference tests are generally permissive if there is much variation in incomes and little variation in (relative) prices. A permissive revealed preference test is unable to produce tight bounds on, for instance, demand in counterfactual price-income regimes.

2. Second, the method implicitly assumes preference stationarity while testing the utility maximisation hypothesis. Therefore, Grüne (2004) argued that a revealed preference test *jointly* verifies consistency with the utility hypothesis *and* stability of preferences. Indeed, preference stability is an inevitable assumption when performing the rationality test. More importantly, this implies that the revealed preference method is necessarily based on repeated observations (of the same individual), because it is even more contestable to assume that preferences are homogeneous across people. The revealed preference approach is therefore less well-suited to deal with cross-sectional data sets.

3. Third, the original axioms (WARP, SARP and GARP) were built on the decisions of a (narrowly defined) homo economicus. Rabin (2002) provided an insightful overview of some of these assumptions. The homo economicus is, for instance, self-interested, purely concerned with final consumption outcomes and discounts his future well-being exponentially. These restrictive assumptions have led many researchers to distrust the revealed preference approach.

There are two ways to deal with Problem 1. The first possibility is to apply the revealed preference principles to experimental data. In an experimental setting, it is possible to control the prices and incomes, and set these elements in such way that discriminatory power is maximised. I refer to Sippel (1997) and Harbaugh et al. (2001) for applications of revealed preference axioms to experimental data. In fact, my contributions in Chapters 1 and 3 also illustrate the usefulness (and strength) of revealed preference tests in an experimental setting. The second possibility is to combine the revealed preference restrictions with nonparametric estimates of Engel curves. This is more common in applications to observational data (for instance expenditure data from budget surveys). Blundell et al. (2003, 2007, 2008) used this procedure to obtain tight bounds on their demand estimates. In Chapter 5, I will present an alternative approach to combine (stochastic) revealed preference restrictions with nonparametric estimates of *population preferences*.

Problem 2 can also be addressed by setting up experiments. In an experiment, the consumption decisions are made in a relatively small time frame, so that preference stability seems a reasonable assumption. Moreover, Stigler and Becker (1977) argue against the abandoning of preference stationarity. It is possible to let the arguments of the preferences change but not the preference itself, or consider a more general (but homogeneous) specification of the preference ordering (cfr supra). Finally, the revealed preference approach of Blundell et al. (2003, 2007, 2008) can be applied to cross-sectional data (provided that Engel curves can be estimated).

Finally, while it is true that many assumptions of the neo-classical approach are empirically unsustainable (Problem 3), it is important to note that the revealed preference framework is flexible enough to characterise alternative assumptions on preferences. Rabin (2002), for instance, discussed an approach to incorporate psychological phenomena into economics, in a way that (still) permits the use of revealed preference theory. Indeed, researchers have made new assumptions about preferences (for example on the arguments in the utility function) without automatically rejecting the assumptions of preference stationarity or utility maximisation. Browning (1989), for example, developed a revealed preference test for consistency with the life cycle model. This model specifically takes into account that consumers care about current *and* future consumption. Moreover, Crawford (2010) presented a revealed preference test of the habits model. The habits model allows the consumer's utility function to depend on past consumption decisions. Demuynck and Verriest (2013) built further on this method to develop a nonparametric test of rational addictions. Summarising, researchers have made considerable progress in incorporating more realistic - psychological - insights into revealed preference models. Chapters 3 and 4 deal with new extensions of the standard revealed preference tests.

I conclude from this that many arguments against the revealed preference literature are, in fact, arguments against a particular theory which is tested or a certain application which is chosen. Hands (2013) correctly argued that

> 'Revealed preference theory is not simply a theory. It is a broad research program in the theory of consumer choice. The revealed preference research program can be thought of as an extended theoretical family - a family containing various family members with different conceptual insights, theoretical structures, and paradigmatic applications.'

General themes in revealed preference This versatility in terms of theoretical and practical applications also implies that the revealed preference method enables researchers to answer a variety of questions. Varian (2006) made an overview of these questions. I will focus on *testing*, *identification* and *prediction*. The thesis will be structured around these main themes.

The first main theme is the *verification* of the rationality hypothesis for a given data set. After all, many economic models are based on the assumption of optimising behaviour on behalf of the economic agents who are analysed. Rationality, albeit crucial to the model and its implications, is taken as given. However, in the spirit of Karl Popper, it is necessary to critically investigate the rationality assumption, and falsify the assumption in settings where it does not hold. Moreover, individuals are heterogeneous in terms of their preferences, their available budget sets and their ability to optimise utility. We can reasonably argue that particular individuals are more likely to be irrational than others. Choosing irrationally has been shown to be equivalent to wasting money. For this reason, it is important to identify who exactly is less rational. Governments could use this information to 'target' their policy to protect the more vulnerable consumers.

The second main theme is the *identification* of the underlying preference structure. The Generalised Axiom of Revealed Preference, for instance, allows to identify the (binary) revealed preference relations from the data. Similarly, the Afriat inequalities allow to recover 'utility levels'. Moreover, Varian (1983) has shown how one can test for additional structure on utility functions, such as weak separability of the utility function in its arguments. Separability is an interesting property from the perspective of *aggregation*. More recently, authors have come up with revealed preference restrictions for collective (household) consumption decisions. The collective model (Chiappori (1988, 1992) and Apps and Rees (1988)) allows household members to have different utility functions, and takes into account that intra-household bargaining power can shift over time. In this framework, recovery can also apply to the identification of the household members' resource shares, which conveys information on their respective bargaining power. Cherchye et al. (2011a) have shown how to use revealed preference techniques to bound the resource shares of both partners, even if the allocation of private goods is unobserved, or in the presence of publicly consumed goods. Information on the sharing rule gives insight into the distribution of welfare in a household.

The third theme is *prediction*. After all, the revealed preference conditions can be applied to bound demand correspondences in counterfactual price-income regimes. Blundell et al. (2003, 2007, 2008) have made seminal contributions. Information on demand responses is relevant both to policy makers and firms. Firms, on the one hand, could use this information to anticipate demand reactions to changing prices. Policy makers, on the other hand, could estimate the impact of policy measures (i.e. that influence the purchasing power of households) on the households' consumption.

Overview of the contributions In my dissertation, I hope to contribute to the *testing*, *recovery* and *prediction* purposes by extending the revealed preference axioms and methods. The extensions either propose a solution to Problems 1 - 3 (see supra) or improve the existing solutions.

In the first part, I deal with *testing consistency* of consumption choices by (individual) children. Given that the setting is experimental, problems of power (Problem 1) and preference stationarity (Problem 2) are mitigated. However, we will argue that the standard consumption experiment - letting children choose from linear budget sets - imposes an overly large computational burden on the respondents. After all, the experiment would require children to exhaust their budget completely. This is not an easy task when more than two goods are presented. The task might be so challenging that it actually interferes with the utility maximisation hypothesis. For this reason, we let children choose from finite choice sets rather than linear budget sets. However, this implies that the standard revealed preference test for individual rationality (i.e. the GARP) must be modified. The standard GARP test is then a sufficient condition for rationality, but not a necessary one. We will come up with alternative rationalisation concepts, depending on whether utility functions are allowed to be monotone, concave, etc.

In the second part, I consider *recovery*. In particular, I want to learn more about the subject of consumers' preferences. I have already argued that the original neo-classical preferences are unrealistic (Problem 3). Several authors have come up with alternative specifications of utility, taking into account future or past consumption. In this part, I consider two alternative extensions of the revealed preference methodology. First, I study other-regarding preferences in the collective framework (see for instance Cherchye et al. (2011a)). I use *experimental* data on joint consumption decisions by children to identify marginal willingness the pay for others' consumption. Second, I introduce diamond goods in a unitary framework. Diamond goods are special in the sense that utility is not only derived from their intrinsic consumption components, or their quantities, but also from their value. I propose a method to capture the 'diamondness' of commodities and apply the approach to *observational* data from the Russian Longitudinal Monitoring Survey. Both papers have in common that a rather restrictive model (unitary or egoistic) is relaxed by introducing an additional parameter (which either captures selfishness or diamondness). In this way, this part also fits in the 'PEEM' (portable extensions of existing models) research program developed by Rabin (2013).

In the third part, I study a method to bound the distribution of welfare measures and demand *predictions* in a setting where panel data are unavailable. In a similar way as Blundell et al. (2003, 2007, 2008), I use nonparametric estimation techniques which, in combination with the revealed preference restrictions, allow for a rather powerful test on the basis of cross-sectional data. This deals with the aforementioned Problems 1 and 2. However, our approach builds on stochastic revealed preference methods. This permits us to include more general forms of unobserved heterogeneity and to bound the distribution of welfare measures and demand correspondences rather than bounding welfare and demand for a representative consumer.

As a final remark, please note that the different chapters in this dissertation are self-contained. The chapters, albeit related, present distinct contributions to the revealed preference literature. For this reason, I modify my notation to the problem under consideration. It is therefore possible that different chapters use different notation. Moreover, I repeat the exposition of GARP at some points in the dissertation, to emphasise the underlying assumptions which are generalised in the respective chapter. Similarly, I discuss the empirical performance measures which are sometimes specifically tailored to one particular data set. In this way, I hope to contribute to a fluent reading of the chapters.

Part II

Revealed preference tests of consistency based on experimental data

In this part, I focus on revealed preference tests of consistency with the utility maximisation hypothesis on the basis of experimental data. I already argued that the revealed preference approach is particularly successful when applied to experimental data. First of all, the assumption of preference stationarity seems more reasonable given the limited time span between the successive decisions. Second, experimenters can choose a particular design (i.e. a set of prices and available budgets) in order to maximise the discriminatory power of the test. Higher discriminatory power implies that the method is better able to reject consistency with the utility maximisation hypothesis when confronted with totally random, irrational choices.

More importantly, an experiment can shed light on consumption decisions which are practically unobservable. The expenditures of children, in particular, are typically not available from budget surveys. However, children influence the consumption behaviour of their parents. They can also spend small amounts of (pocket) money. In this sense, their decisions matter for economic analysis. Furthermore, insight into the rationality of children is important from a 'paternalist' point of view. Parents should protect children who are vulnerable to irrational decisions.

I apply the revealed preference approach to (experimental) consumption decisions by individual children. The analysis is similar in spirit to the experiment set out by Harbaugh et al. (2001). In the same way, children choose commodity bundles from finite choice sets. It seems reasonable to argue that choosing from finite choice sets is less cumbersome than spending a given amount of money on various goods with different prices. Budget exhaustion would require the children to use calculators (especially in a setting with more than two goods). The computational burden would likely interfere with the optimality of children's choices. However, the aim of this part is to obtain insight into the rationality of children and the drivers of rationality.

In Chapter 1, I investigate whether rationality is related to intelligence. Harbaugh et al. (2001) briefly touched upon the subject, by linking rationality to mathematical skills. In

Chapter 1, I use teacher based assessments of intelligence to investigate whether, and to what extent, smart children tend to behave more rational. I specifically recognise the multidimensional nature of intelligence. More generally, this chapter illustrates consistency checks for experimental data as well as a number of common empirical performance measures in the revealed preference literature, such as pass rates, power and the violation index.

In Chapter 2, I present revealed preference conditions for consistency with utility maximisation in a finite choice-set setting. Finite choice sets are explicit in the experiments in Chapter 1 and Harbaugh et al. (2001), but they also occur in many real-life settings. When choice sets are finite, it is necessary to make a distinction between the cases where the underlying utility function is weakly monotone, strongly monotone and/or concave. Moreover, I discuss a number of conditions under which the usual revealed preference test (i.e. GARP) is still valid. Finally, I compare results from the different rationalisability concepts with the results from Chapter 1 and the results from Harbaugh et al. (2001), who tested for consistency with a strongly monotone utility function. The results clearly depend on which rationalisability concept is applied.

Chapter 1

Are the smart kids more rational?[1]

1.1 Introduction

We use experimental data to study the 'rational' consumption behaviour of children. Considering children of different ages, we assess the empirical validity of the rationality assumption. Next, we also explain the degree of rationality in terms of the children's personal characteristics. In this respect, a specific feature of our study is that we relate rational consumption to alternative dimensions of intelligence. In particular, we investigate how verbal skills (language) and non-verbal skills (mathematics) define the (ir)rational consumption behaviour of children. Or putting it differently, are the 'smart' kids more rational? And, if so, does the type of smartness matter? This introductory section motivates our research question, and indicates how this study relates to the existing literature.

Motivation. The literature has devoted considerable attention to studying whether economic models are applicable to children.[2] The aim is to understand the children's decision

[1] This chapter is based on joint work with Sabrina Bruyneel (KU Leuven), Laurens Cherchye (KU Leuven), Bram De Rock (ULB) and Siegfried Dewitte (KU Leuven). I refer to the working paper version of Bruyneel et al. (2012a).

[2] See, for example, Harbaugh et al. (2001, 2002, 2003, 2007), Farrell and Shields (2007), Lundberg et al. (2009), and references therein.

behaviour, and to gain insight into the evolution of this behaviour when children grow older. Clearly, a better understanding of children's economic behaviour allows for a better modelling of this behaviour. For instance, household consumption models that include children usually treat these children as either some 'public good' entering their parents' utility functions or as rational decision makers maximising their own utility.[3] This immediately raises the question whether and to what extent children can actually be considered as rational consumers. We refer to Harbaugh et al. (2001) and Lundberg et al. (2009) for a more elaborated discussion of this type of arguments.

Next, there is the obvious observation that children often do have to make consumption decisions (e.g. on how to spend their pocket money or how to choose the games they play). In this respect, Choi et al. (2014) argue, rather convincingly, that the quality of such consumption choices can be measured by the degree of rationality. In particular, they show that irrational choices imply a 'waste of money'. In a similar vein, Echenique et al. (2011) indicate that irrational consumers are subject to being exploited as a 'money pump'. This pleads for protecting and guiding children's behaviour more carefully if this behaviour turns out to be irrational. In this respect, identifying the children's characteristics that drive rationality can also lead to more effective protection and training of those children who are particularly vulnerable.[4]

Methodology and related literature. Observational 'real-life' data usually do not contain sufficient information on children's consumption to study rationality of their behaviour. Therefore, most papers cited above make use of experimental data to study children's behaviour.[5] In line with this common practice, we also conduct an experiment in the current

[3]For example, Blundell et al. (2005) and Cherchye et al. (2012) propose household consumption models that treat children as public goods, while Becker (1974), Dauphin et al. (2011) and Dunbar et al. (2013) consider models that assume children are individually rational decision makers.

[4]For example, Choi et al. (2014) argue that insight into the relationship between rationality and personal characteristics can prove useful to design appropriate social programs (Manski (2001)) and paternalistic policies (Thaler and Sunstein (2003)).

[5]One exception is the study of Farrell and Shields (2007), which uses child diary information contained in the British Family Expenditure Survey. However, the cross-sectional data used by these authors do not allow for testing

study. Specifically, for each individual child we gather data on 9 (unsophisticated) discrete consumption choices. As we will indicate, this experimental setup effectively allows for a powerful test of rationality.

Our experimental design is similar to the one of Harbaugh et al. (2001), who also focused on verifying whether children can be modelled as rational decision makers. However, it is worth emphasising two main differences between these authors' study and ours, which -in our opinion- make our study a valuable extension of this original one. Firstly, as for our testing methodology, we use a different procedure to check rationality, which is specially tailored for discrete choice settings such as ours. Secondly, and more importantly, our experimental data set is richer than the one of Harbaugh, Krause and Berry in that it includes more detailed information on the children's personal characteristics. Most notably, as indicated above, we have information on alternative dimensions of child intelligence. As we will explain, this effectively provides a more balanced insight into the driving forces of economically rational behaviour. It will turn out to be important to explicitly account for the multidimensional nature of intelligence to identify significant effects.[6]

At the methodological level, we use so-called 'revealed preference' tests to check rationality of the children's consumption behaviour.[7] A particular feature of this revealed preference approach is that it starts directly from the observed choices and does not require any functional specification of the individual preferences. It directly verifies the testable implications of rationality on the raw consumption data. Conveniently, this avoids that rationality of observed behaviour is rejected simply because of functional misspecification. Moreover, revealed preference tests can be meaningfully applied to small data sets. For our experiment, this means that we can analyse rationality of choices for each child separately and, thus, that

the rationality of the children's decision behaviour itself.

[6]In this respect, Harbaugh et al. (2001) only considered the effect of mathematical skills in their original study, and their results did not reveal a significant relation between rationality and mathematical ability. In Section 1.5, we will indicate that the effect of mathematical ability on rationality only appears significantly if one controls for differences in verbal skills (in our case language). This shows the importance to simultaneously account for multiple dimensions of intelligence when studying its effect on rationality of consumption behaviour.

[7]See Samuelson (1938), Houthakker (1950), Richter (1966), Afriat (1967), Diewert (1973) and Varian (1982) for seminal contributions on the revealed preference approach that we adopt here.

we can avoid the debatable assumption that (e.g. observably similar) children have homogeneous preferences. Finally, and also because of these reasons, it has been argued in the literature that revealed preference tests are specially useful within an experimental context such as ours.[8]

Paper outline. Section 1.2 describes the specificities of our experiment. As indicated above, we confronted each child with 9 choice problems. In each problem, the child had to choose between 7 commodity bundles, which makes that we observe 9 consumption choices in total. In this section, we will also explain how we obtained our variables on the children's personal characteristics. In particular, following the argumentation of Hoge and Coladarci (1989) we use teacher assessments to construct our indicators of intellectual skills.

Section 1.3 presents our revealed preference test of rationality. Because our experiment involves discrete choices, we cannot straightforwardly apply the usual revealed preference tests, which are designed for standard (continuous) budget sets.[9] Therefore, we discuss an adapted revealed preference test that can deal with our type of discrete choice data.

Section 1.4 presents the results of our revealed preference tests for the children in our sample. We discuss pass rates and discriminatory power of our test. In addition, if a child's observed consumption behaviour turns out to be irrational, we measure how close it is to rationality by using the violation index.[10] Intuitively, this index evaluates the degree of rationality by the amount of money that is wasted by making irrational decisions. Adopting the argument of Choi et al. (2014) that we cited above, this index thus quantifies the 'quality' of the observed consumption decisions.

Section 1.5 investigates the personal characteristics that drive rational behaviour. A particular focus is on the question whether and to what extent 'being smart' relates to 'being

[8]See, for example, Sippel (1997), Harbaugh et al. (2001), Andreoni and Miller (2002), Caplin et al. (2011), Bruyneel et al. (2012b) and Caplin and Dean (2014). Cox (1997) also provides an extensive discussion on the use of revealed preference methodology in combination with experimental data.

[9]See, for example, Varian (1982) for more discussion on standard revealed preference tests.

[10]See Varian (1993) for a discussion of this violation index.

rational'. A main finding here will be that such an investigation must take the multidimensional nature of intelligence (verbal versus non-verbal skills) into account.

Section 1.6 concludes.

1.2 The experimental data

In this section, we first describe how we obtained our data on children's personal characteristics. Next, we explain our experiment to collect (discrete) choice data. We will also indicate how this experimental design impacts on the revealed preference test that we present in Section 1.3.

Personal characteristics. Our sample includes a total of 100 children residing at three different schools that participated to our experiment (39 from kindergarten, 31 third graders and 30 sixth graders). The selection of classes and schools in the sample is presented in Table 1.1.

	kindergarten	third grade	sixth grade
School I	1 class (15)	1 class (16)	1 class (11)
School II	1 class (13)	1 class (15)	1 class (19)
School III	1 class (11)	0 classes	0 classes

Table 1.1: Information on schools and classes (number of children per class)

Child ages range from 5 to 12 years, with a mean value of approximately 8 years. Hoge and Coladarci (1989) argue that teacher based assessments form a reliable source of information on children's characteristics[11] because teacher based assessments and achievement test scores are typically highly correlated. Following the argumentation of Hoge and Coladarci (1989), we asked teachers about each child's intellectual skills, which include language as a verbal skill and mathematical ability as a non-verbal skill. We also consider creativity as an

[11] In fact, Borkenau and Liebler (1993) argue that acquaintances can, quite accurately, assess a person's intelligence. In a recent paper by Lonnqvist et al. (2012), for instance, teacher and parent ratings are used to evaluate the cognitive abilities of children.

additional dimension of intelligence[12]. Next, we also asked for the number of older siblings in the child's family. We will motivate our use of data on older siblings in Section 1.5.

We used two different indicators to quantify the three intellectual skills. The first indicator has possible scores ranging from 1 (= bottom 2 % compared to peers) to 8 (= top 2 % compared to peers), and the second indicator has possible scores ranging from 1 (= very weak compared to peers) to 10 (= very strong compared to peers). See Appendix 1.A for more details. Our following empirical exercises will make use of a composite of these two indicators[13], which is constructed in two steps. First, we transformed (i.e. multiplied by 10/8) the scores for the first indicator so that they also ranged from 1 to 10. Subsequently, our composite skills indicators are computed as the average of these transformed scores for the first indicator and the original scores for the second indicator. As an alternative we can select only one indicator per dimension of intelligence. This does not affect the results in Section 1.5 too much because of the high correlation between both indicators.

Table 1.2 provides summary information on the children's personal characteristics for our sample. Generally, for the different characteristics we obtain quite some variation across the participants of our experiment. This observation is particularly useful in view of our explanatory analysis in Section 1.5, where we will relate this variation to differences in rationality of the children's consumption behaviour.

	obs	mean	std dev	min	max
age	100	8.04	2.825	5	12
mathematics	99	7.364	1.686	2.75	10
language	99	7.283	1.601	2.75	10
creative	99	7.24	1.489	3.875	10
older siblings	93	0.935	0.987	0	5

Table 1.2: Summary statistics for children's characteristics

[12] The triarchic theory of intelligence (Sternberg (1985)), for instance, stresses the importance of creativity.

[13] These indicators are highly correlated, with correlations ranging from 0.899 (for the creativity measures) to 0.925 (for the mathematics measures).

Experimental design. As indicated in the Introduction, our experiment is similar in design to the one of Harbaugh et al. (2001). As a starting point, we take from these authors that it does not seem appropriate to ask the participants of our experiment to select commodity bundles from a continuous budget set (defined by given prices and budget). Because such a selection process involves abstract mathematical reasoning, this seems too difficult a task for young aged children. The complexity of the choice problem - the children do not understand it - and its computational difficulties might interfere with the desire to make optimal decisions.

To account for this difficulty, we confronted the children in our experiment with 9 unsophisticated discrete choice problems, each characterised by a choice set C_t consisting of (only) 7 consumption bundles. More precisely, we conceived 9 different price regimes. For a fixed budget, each such price regime defines a budget hyperplane. Then, we selected 7 quantity bundles from every budget hyperplane, and in each choice problem we asked the children to pick their preferred bundle from these 7 bundles. Thus, our experiment is such that children did not face *explicit* prices and budgets when selecting their consumption bundles, but we can interpret their choices as defined under *implicit* prices and budgets. Clearly, this considerably facilitated the children's decision process; they simply had to select 9 commodity bundles from choice sets C_t that contained only 7 elements.

More specifically, the children could choose quantity bundles that were composed of three commodities: grapes (units of 10 grams), mandarins (units of 12.5 grams) and letter biscuits (units of 5 grams). The 7 bundles in a given choice set always contained the three 'extreme' bundles with all budget spent on a single commodity, and four other 'intermediate' bundles with positive quantities for our three commodities. Appendix 1.B reports the choice sets we used and the associated (implicit) prices[14]. In Section 1.4, we will argue that these choice sets imply a powerful revealed preference test of rationality.

[14] Although some bundles included fractions of the commodities, the experimenter did not report aversion to these bundles. The children were provided with separate segments of the mandarin. Grapes and letter biscuits were distributed in units.

The whole experiment is carried out in the classrooms of the four participating schools. Each child is invited to taste the grapes, mandarins and letter biscuits prior to the experiment. These tasted commodities are of the same brand and have the same quality as the commodities in the actual decision problems. Finally, the children were truthfully told that, at the end of the experiment, they would receive one consumption bundle that would be taken randomly from the 9 bundles they selected. This should improve the external validity of our study. We provide more details on our experiment in Appendix 1.B. Importantly, given our specific setup, we may safely abstract from intertemporal issues like savings and interdependent consumption choices.

Table 1.3 provides summary information on the individual budget shares of the three different commodities. The fact that children chose positive amounts of all commodities makes us conclude that each commodity is effectively desirable. On average, the letter biscuits commodity seems to be the more popular one. However, the reported standard deviations also reveal quite some heterogeneity over the choices made. Generally, all this suggests that our experimental data provide a useful basis to assess rationality of consumption choices.

Table 1.3 also gives separate budget share information for female and male participants in our experiment. It seems that there are no specific gender effects in terms of the commodities that are selected. Actually, this is a systematic finding for our sample of children: for none of the exercises discussed in Sections 1.4 and 1.5, we found significant differences between male and female participants. Therefore, and to compactify our exposition, we will not report on gender effects in these following exercises. Evidently, results for the male and female subsamples are available from the authors upon simple request.

As a final point, it is worth noting that our discrete choice setting raises some particular issues. Firstly, it could well be that for a number of children the 7 options in some choice problem did not include the children's most preferred bundle defined over the entire budget set associated with the corresponding (implicit) prices and budget. However, during our experiment it was clear that our selection of choices strikes a right balance between simpli-

	grapes	mandarins	letter biscuits
all children	.317	.245	.438
	(.172)	(.190)	(.258)
kindergarten	.273	.174	.553
	(.204)	(.180)	(.284)
third grade	.312	.235	.452
	(.117)	(.173)	(.208)
sixth grade	.378	.347	.275
	(.159)	(.179)	(.178)
female	.307	.245	.448
	(.170)	(.184)	(.263)
male	.329	.244	.427
	(.175)	(.199)	(.255)

Table 1.3: Average budget shares (standard deviations between brackets)

fying the problem and exhausting all budget possibilities (see also our empirical analysis in Section 1.4). Secondly, and related to this, standard revealed preference tests assume continuous budget sets, while our setup implies discrete choice sets. This makes that we need an adaptation of the standard revealed preference test of rationality, which we discuss next.

1.3 A revealed preference test of rationality

As explained in the previous section, the children were faced with 9 different discrete choice sets C_t. Each choice set contained 7 quantity bundles taken from a budget hyperplane.[15] The children had to choose one quantity bundle $\mathbf{q}_t \in \mathbb{R}^3_+$ (with $t = 1, \ldots, 9$) from this set. For each individual child, this defines the data set

$$S = \{(C_t; \mathbf{q}_t), t = 1, ..., 9\}.$$

The data set S is consistent with rationality if there exists a continuous and strongly monotone utility function U which rationalises the data, in the following sense.

[15] More precisely, there exist (implicit) prices $\mathbf{p}_t$ and an (implicit) budget y_t such that, for all $\mathbf{z} \in C_t$, $\mathbf{p}_t \mathbf{z} = y_t$.

Definition 1.1. Let $S = \{(C_t; \mathbf{q}_t); t = 1, ..., 9\}$ be a set of observations. A strongly monotone utility function U provides a *rationalisation* of S if and only if for each observation $t = 1, ..., 9$ we have $U(\mathbf{q}_t) \geq U(\mathbf{z})$ for all $\mathbf{z} \in C_t$.

Note that, although we used implicit prices and budgets to design the experiment (i.e. to describe the hyperplanes used in the construction of the choice sets C_t), we do not make use of this information in our definition of rationality and, thus, we cannot use this information in the corresponding revealed preference test of rationality. As an implication, our test results will not depend on the fact that the children were not aware of the actual price-budget information.

The following concepts will be crucial ingredients of our revealed preference test of rationality.

Definition 1.2. Let $S = \{(C_t; \mathbf{q}_t); t = 1, \ldots, 9\}$ be a set of observations. Then for any $s, t = 1, \ldots, 9$:

(i) $\mathbf{q}_t$ is directly revealed preferred over $\mathbf{q}_s$ (i.e. $\mathbf{q}_t R_0 \mathbf{q}_s$) if there exists a $\mathbf{z} \in C_t$ such that $\mathbf{z} \geq \mathbf{q}_s$;

(ii) $\mathbf{q}_t$ is strictly directly revealed preferred over $\mathbf{q}_s$ (i.e. $\mathbf{q}_t P_0 \mathbf{q}_s$) if $\mathbf{q}_t R_0 \mathbf{q}_s$ and $\mathbf{q}_s \notin C_t$;

(iii) $\mathbf{q}_t$ is revealed preferred over $\mathbf{q}_s$ (i.e. $\mathbf{q}_t R \mathbf{q}_s$), if there exist observations $u, v \ldots, w$ such that $\mathbf{q}_t R_0 \mathbf{q}_u$, $\mathbf{q}_u R_0 \mathbf{q}_v$, ..., $\mathbf{q}_w R_0 \mathbf{q}_s$.

Essentially, this definition makes clear the preference information we can extract from a child's observed choices contained in some data set S. The intuition goes as follows. As for statement (i), if the child chooses $\mathbf{q}_t$ in C_t, then (s)he must prefer $\mathbf{q}_t$ over all available bundles $\mathbf{z} \in C_t$. Because we assume strictly increasing utility functions (i.e. strongly monotonic preferences), this also means that the child prefers $\mathbf{q}_t$ over any $\mathbf{q}_s$ for which $\mathbf{z} \geq \mathbf{q}_s$. Statement (ii) builds further on the first statement and concludes from $\mathbf{q}_t R_0 \mathbf{q}_s$ that there exists $\mathbf{z} \in C_t$ such that $\mathbf{z} \geq \mathbf{q}_s$. Then, because $\mathbf{q}_s \notin C_t$ we must have that $\mathbf{z} \neq \mathbf{q}_s$, which means that $\mathbf{z}$ strictly dominates $\mathbf{q}_s$ in at least one commodity while not having less of any other commodity. Strictly increasing utility then implies that $\mathbf{z}$ is strictly preferred to $\mathbf{q}_s$, which carries over to $\mathbf{q}_t$ being strictly preferred to $\mathbf{q}_s$. Finally, statement (iii) imposes that preferences are transitive.

While Definition 1.2 resembles the standard definition of revealed preference relations (as, for example, in Varian (1982)), it is substantively different because it is defined in terms of discrete choice sets and does not use price information to reconstruct the preference relations. More precisely, let $\mathbf{p}_t$ and y_t represent the prices and budget information that describes the choice set C_t (i.e. for all $\mathbf{z} \in C_t : \mathbf{p}_t'\mathbf{z} = y_t$). Then, it is easy to verify that $\mathbf{q}_t R_0 \mathbf{q_s}$ (respectively $\mathbf{q}_t P_0 \mathbf{q_s}$) implies that $\mathbf{p}_t'\mathbf{q}_t \geq \mathbf{p}_t'\mathbf{q}_s$ (respectively $\mathbf{p}_t'\mathbf{q}_t > \mathbf{p}_t'\mathbf{q}_s$). Importantly, however, the reverse implication does not hold necessarily, because the discrete choice set C_t does not contain all the bundles on the continuous budget hyperplane defined by $\mathbf{p}_t$ and y_t (i.e. the set $\{\mathbf{z} : \mathbf{p}_t'\mathbf{z} = y_t\}$). Consider for instance the data set presented in Figure 1.1.

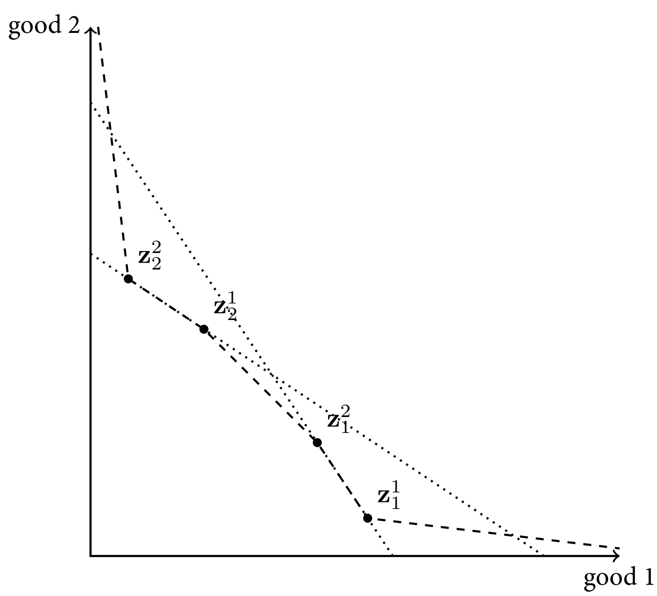


Figure 1.1: Rational choices that violate GARP

There are two choice sets $C_1 = \{\mathbf{z}_1^1, \mathbf{z}_1^2\}$ and $C_2 = \{\mathbf{z}_2^1, \mathbf{z}_2^2\}$. We assume that $\mathbf{z}_1^1$ is chosen from C_1 ($\mathbf{q}_1 = \mathbf{z}_1^1$) and $\mathbf{z}_2^2$ is chosen from C_2 ($\mathbf{q}_2 = \mathbf{z}_2^2$). It is easy to see that these choices violate GARP. However, the choices are rationalisable by a strongly monotone and non-concave utility function if we take the finiteness of the choice sets (C_1 and C_2) into account. For example, the dashed curve provides a (non convex) indifference curve which rationalises all choices.

The following proposition defines a revealed preference test of rationality for our discrete choice setting. In particular, it provides a necessary and sufficient condition for a strongly monotone rationalisation of the information contained in the data set S. Chapter 2 (specifically: Appendix 2.A, Theorem 2.4, proof for SMARP) contains the formal proof of this result.[16]

Proposition 1.3. Let $S = \{(C_t; \mathbf{q}_t)\,; t = 1, \ldots, 9\}$ be a set of observations. Then, there exists a strongly monotone utility function U that provides a *rationalisation* of S if and only if, for any $s, t = 1, \ldots, 9$, $\mathbf{q}_t R \mathbf{q}_s$ implies not $\mathbf{q}_s P_0 \mathbf{q}_t$.

The revealed preference condition in this result states that, if some bundle $\mathbf{q}_t$ is revealed preferred over a bundle $\mathbf{q}_s$, then it cannot be that $\mathbf{q}_s$ is strictly directly revealed preferred over $\mathbf{q}_t$. In other words, we cannot have a cycle of revealed preference relations containing a direct revealed preference relation that is strict. The rationality condition in Proposition 1.3 is closely related to the Generalised Axiom of Revealed Preference (*GARP*; see again Varian (1982)), which is the standard condition for rationality in the revealed preference literature. Similar to before, the important difference is that our rationality condition is defined for discrete choice sets and, unlike *GARP*, does not use price information. As a final remark, it is important to note that different rationalisation concepts, which would coincide in a setting with linear budget sets, no longer coincide when choice sets are finite. Specifically,

[16] As is clear from the text, the following rationality condition is specific to discrete choice settings in which each choice set contains a finite number of bundles taken from one and the same budget hyperplane. We can refer to Harbaugh et al. (2001), Polisson and Quah (2013) and Chapter 2 for revealed preference tests of rationality in alternative discrete choice settings.

different assumptions on the monotonicity and/or concavity of the utility function can lead to different results. This point is further elaborated in Chapter 2.

1.4 Testing rationality

In this section, we present our test results for the children under study. We begin by considering pass rates for our rationality tests. Here, a particular focus will be on whether these pass rates vary depending on age. Next, to enable a better interpretation of these pass rates, we also compute the discriminatory power of the revealed preference test for our experimental design. In the current context, power stands for the probability of detecting (simulated) irrational behaviour. Finally, we also report results on the violation index for our sample. As indicated in the Introduction, this measure quantifies how close observed behaviour is to rational behaviour, which actually allows us to measure the 'quality' degree of the observed consumption decisions.

Pass rates. As mentioned before, our setup allows us to carry out the revealed preference test for each child separately. Per respondent, we obtain a positive response if his or her choices can be rationalised and a negative response if his or her choices are not rationalisable. As such, we obtain 100 independent tests of rationality. Pass rates then capture the average response across the sample. Table 1.4 presents the results for each age category. We learn that pass rates are generally low but increasing with age. However, even the older children in our sample behave irrationally.

To put this finding into perspective, it is useful to compare the results in Table 1.4 with the ones obtained by Bruyneel et al. (2012b). For a similar experiment (with three goods and the same prices and budgets) on undergraduate students, these authors obtained a substantially higher pass rate of 92%. As such, we can conclude that younger aged children appear considerably less rational than young adults.

At this point, we remark that this conclusion is partly at odds with the results of Harbaugh et al. (2001). These authors focused on children of the second and sixth grade, and compared their results for these groups to the ones for undergraduate students. For the second graders they also obtained a very low pass rate (of 26%). However, the pass rate for the sixth graders was closely similar to the one for the undergraduate students (i.e. both rates were situated between 60% and 65%). One possible explanation for our different results is that we consider choice problems involving three goods, whereas Harbaugh, Krause and Berry concentrated on a simpler setting with only two goods. One may argue that adding goods makes consumption decisions more difficult, and that this effect is more pronounced for younger consumers.

	pass rate	obs
all children	0.43	100
	[0.331;0.529]	
kindergarten	0.31	39
	[0.162;0.458]	
third grade	0.48	31
	[0.301;0.659]	
sixth grade	0.53	30
	[0.348;0.712]	

Table 1.4: Individual rationality: pass rates, [95 per cent confidence bounds]

Power. To enable a better interpretation of the pass rates in Table 1.4, we also computed the discriminatory power of our rationality test. This power is defined as the probability of detecting 'irrational' behaviour (i.e. behaviour that is not consistent with utility maximisation as characterised in Definition 1.1). As such, this power value provides a natural benchmark for the pass rates that we discussed above: if power is situated below the rejection rates for our sample (i.e. one minus the pass rate), then we can conclude that observed behaviour appears *even less* rational than (simulated) irrational behaviour, which obviously provides a strong rejection of the rationality hypothesis.[17]

To simulate irrational behaviour, we use the bootstrap method for panel data as described by Andreoni and Miller (2002) within a similar experimental context.[18] Essentially, this method mimics random behaviour for each choice set by drawing randomly from the children's observed choices for that set (i.e. 100 choices for each of the 9 different choice sets in our setting). In other words, the bootstrap procedure draws bundles from the observed probability density function. This gives insight into the expected distribution of violations under random choice, while incorporating information on the observed choices. Corresponding results for an alternative power method (Bronars' approach) are discussed in Chapter 2.

More specifically, we conducted Monte Carlo-type simulations that include 10000 iterations. This obtained a power value for our test of 0.87. In other words, the null hypothesis of rational - optimising - behaviour is rejected with a probability of 87% when the alternative hypothesis holds. The alternative hypothesis stipulates, in this case, that the observed choices in the sample constitute a set and that choices are made from this set at random (with replacement). A power measure of 0.87 is clearly very high, which confirms that the experiment is well-designed. Also, and importantly, it is much above the rejection rates that can be computed from Table 1.4. In our opinion, this provides some (albeit moderate) support in favor of the null hypothesis of optimising behaviour for our sample of children: even though our pass rates are rather low in absolute terms, they are reasonably high in view of the power of our test.

For robustness, we also consider an alternative bootstrapping procedure which takes into account that the distribution of choices is specific to the age group under consideration. In particular, we drew (random) bundles from the observed distribution of choices by, respectively, kindergarten respondents, third graders and sixth graders. This gives a

[17] See Beatty and Crawford (2011) for an extensive discussion of this interpretation.

[18] We refer to Bronars (1987) and Andreoni et al. (2011) for more discussion on alternative methods to measure the discriminatory power of revealed preference tests.

bootstrap power of 0.91 for kindergarten respondents, 0.78 for third graders and 0.74 for sixth graders. Again, these power estimates are considerably higher than the corresponding rejection rates in Table 1.4.

Violation index. As a further investigation, we also computed a *violation index* θ for the choices in our experiment. The violation index relaxes the expenditure minimisation requirement conditional on a set of revealed preference relations. To compute this index, we make use of the (implicit) prices and budgets underlying the construction of our choice sets. More precisely, for each choice observation t we compute

$$\theta_t = \frac{\min_{\mathbf{q}_s R \mathbf{q}_t} \mathbf{p}_t' \mathbf{q}_s}{\mathbf{p}_t' \mathbf{q}_t}.$$

This violation index takes the finiteness of the choice sets into account when constructing revealed preference relations R. Otherwise stated, R refers to the corresponding relation in Definition 1.2.

Consider for instance Figure 1.2 which presents two choice sets $C_1 = \{\mathbf{z}_1^1, \mathbf{z}_1^2\}$ and $C_2 = \{\mathbf{z}_2^1, \mathbf{z}_2^2\}$. Suppose that the implicit prices corresponding to the choice sets are given by

$$\begin{aligned} \mathbf{p}_1 &= [3, 2] \\ \mathbf{p}_2 &= [2, 3]. \end{aligned}$$

A respondent has chosen bundle $\mathbf{q}_1 = \mathbf{z}_1^1$ from choice set 1 and bundle $\mathbf{q}_2 = \mathbf{z}_2^2$ from choice set 2.

$$\begin{aligned} \mathbf{q}_1 &= [3, 3/2]' \\ \mathbf{q}_2 &= [3/2, 3]'. \end{aligned}$$

One can see that this data set is rationalisable (in a discrete setting), and specifically, $\mathbf{q}_2$ R $\mathbf{q}_1$. However, it is possible that $\theta_1 < 1$, indicating a waste of the implicit budget. In a non-discrete setting, the respondent could have saved money (specifically, $\mathbf{p}_1(\mathbf{q}_1 - \mathbf{q}_2) = 3 \cdot (3 - 3/2) + 2 \cdot (3/2 - 3) = 1.5$) in period 1 by purchasing $\mathbf{q}_2$ instead of $\mathbf{q}_1$ because $\mathbf{q}_2$ is also preferred over $\mathbf{q}_1$. Graphically, the dashed line presents the lower expenditures in period 1 associated with bundle $\mathbf{q}_2$. However, $\mathbf{q}_2$ was not included in the discrete choice set in period 1. The discreteness of the choice set thus leads to an implicit money waste although the respondent is fully rational. For this reason, θ_t gives the extent to which the consumer is vulnerable to money losses. This vulnerability depends on the individual's preferences, his or her level of rationality and the construction of the choice set. Therefore, θ_t (i.e. the amount of money wasted) gives a lower bound on rationality. The condition that $\theta_t = 1$ is sufficient to impose (exact) rationality. By contrast, lower values indicate that part of the implicit budget may have been wasted.

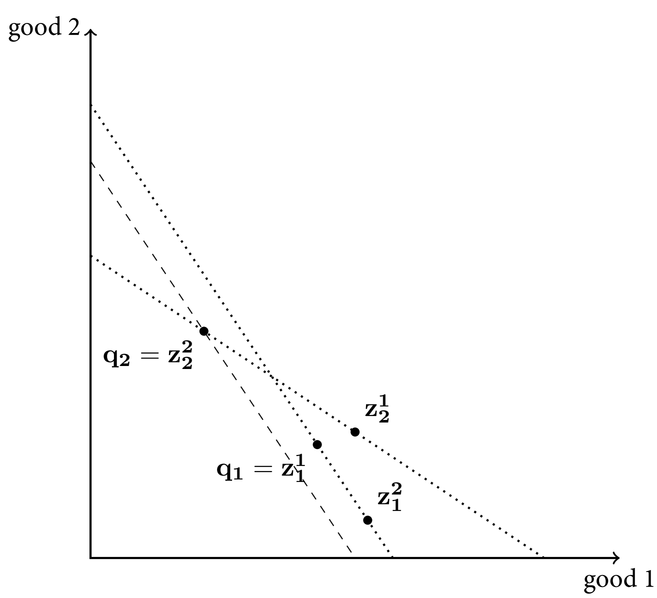


Figure 1.2: Illustration of the violation index in a finite choice set setting

The violation index θ is then simply the minimum θ_t defined over all t:

$$\theta = \min\{\theta_1, ..., \theta_9\}.$$

Building further on the above explanation of θ_t, Choi et al. (2014) interpret this violation index as a measure of the quality of the observed consumption decisions.[19]

Table 1.5 presents some descriptive statistics for the violation index. We find that the mean value of the index equals 69%, which means that children 'waste' no more than 31% of their budget, on average. Similar to our results on pass rates, this can be interpreted as signalling a rather low quality of the children's decision making process. At this point, however, it also worth remarking that there appears to be quite some heterogeneity across individual children. For example, just like for our pass rates, we again observe an age effect: third graders and sixth graders waste no more than 27% of their budget. This is about 10% to 15% less than the corresponding budget waste by children from kindergarten. Finally, we also computed the violation index for random data simulated along the lines described in the previous paragraph. We obtain a violation index of 43% for the random data. This clearly indicates that the children's behaviour - albeit not fully rational - is more consistent with the utility maximisation hypothesis than the random, simulated data sets. In the next section, we provide a more in-depth analysis of (observable) characteristics impacting on the quality of children's consumption decisions.

Variable	obs	mean	std dev	min	max
all children	100	.688	.361	.11	1
kindergarten	39	.604	.382	.111	1
third grade	31	.737	.321	.11	1
sixth grade	30	.747	.363	.111	1

Table 1.5: Individual rationality: a violation index

[19] See Varian (1993) for more discussion on the violation index.

1.5 Explaining rationality

As a preliminary step, we consider Table 1.6, which provides descriptive statistics on the characteristics of both the group of rational children (i.e. children that pass our revealed preference test) and the group of irrational children (i.e. children that fail our revealed preference test). We again observe an age effect. Moreover, there seem to be some differences for the personal characteristics. However, these differences do not appear to be very pronounced. We find no significant differences when comparing mathematical or language skills for the two groups of children under consideration. This finding falls in line with the one of Harbaugh et al. (2001), who also considered the relationship between rationality and mathematical ability, and did not detect a significant effect either.

Variable	nr rat	nr irrat	mean rat	mean irrat	sd rat	sd irrat
age**	43	57	8.605	7.614	2.77	2.814
mathematics	43	56	7.358	7.368	1.855	1.561
language	43	56	7.061	7.453	1.713	1.503
creative*	43	56	7.439	7.087	1.604	1.39
older siblings	40	53	.925	.943	.944	1.027

Table 1.6: Characteristics of rational and irrational children (* = significant difference at 10%; ** = significant difference at 5%)

Importantly, however, the mean values in Table 1.6 are unconditional in nature. For example, computing the average difference of mathematical skills for rational and irrational children does not correct for language differences between these two groups of children. To allow for a conditional analysis, we will next regress our rationality results simultaneously on alternative dimensions of intelligence. This will lead to more refined insights because it effectively exploits the multidimensional nature of intelligence[20]. In fact, this also makes better use of the richness of our data on personal characteristics.

Following up on our discussion in the Section 1.4, we conduct two types of regression exercises: probit regressions with the rationality indicator (0 if the child is irrational and 1

[20]Following the well-known Cattell-Horn-Carroll theory (Carroll (1993)), IQ must be considered as a multidimensional concept.

if the child is rational) as dependent variable and fractional response regressions with the violation index (situated between 0 and 1) as dependent variable. To estimate the latter type of regression, I follow a procedure set out by Papke and Wooldridge (1996). These authors proposed quasi-maximum likelihood to estimate a fractional response model when the dependent variable is bounded between 0 and 1, with many observations at the boundary (in casu: 1)[21]. While a standard regression (OLS) on transformed data (e.g. the logit transformation) cannot properly deal with values at the boundary, the enhanced generalised linear model proposed by Papke and Wooldridge (1996) can take these values into account.

In each regression, we include age dummies (kindergarten and third grade) and the number of older siblings as control variables. Given our previous results we expect negative signs for the age dummies (especially for the kindergarten coefficient). Next, the use of older siblings as a control is inspired by the literature claiming that living in a family with older siblings impacts positively on the rationality of the younger children, because it makes these younger children better informed decision makers.[22]

Our core focus is then on investigating how verbal skills (language) and non-verbal skills (mathematics) define rationality of children's behaviour. In particular, we address two questions. First, we want to verify whether the smarter kids are more rational. In this respect, it is worth to recall that our four intellectual characteristics are measured by comparing the children to their peers. As such, if a child achieves a high score for a particular dimension, this indicates that this child does relatively well as compared to other children of her/his age. Next, we also aim to identify the specific personal characteristics that drive rationality. Is it the case that some intellectual characteristics bear a stronger relation to rationality than others? And does the effect of alternative dimensions of intelligence move in opposite directions? A priori, because rational consumption behaviour may be argued to require abstract/mathematical thinking, one may believe that mathematical ability will relate positively

[21] This method requires specifying robust standard errors, a logit link function and the binomial distribution. It is available in STATA under the command 'glm y x, link(logit) family(binomial) robust'.

[22] See, for example, John (1999) for a recent survey.

to rationality, while the expected effect of language may be less pronounced.

Table 1.7 presents our regression results. For our two dependent variables (i.e. the rationality indicator and the violation index) we have analysed two regression specifications. Our main focus will be on specification 1, which (only) includes language and mathematical skills as indicators of (verbal and non-verbal) intellectual ability. As a robustness check, we also considered specification 2, which includes creativity as an additional dimension of intelligence.

	rationality indicator (0 or 1)		violation index (between 0 and 1)	
	specification 1	specification 2	specification 1	specification 2
kindergarten	-.588*	-.564*	-.677	-.631
	(.331)	(.339)	(.456)	(.463)
third grade	-.028	.113	-.042	.042
	(.345)	(.352)	(.471)	(.465)
mathematics	.247*	.265*	.293*	.291*
	(.138)	(.158)	(.169)	(.175)
language	-.321**	-.506***	-.382**	-.475**
	(.14)	(.175)	(.179)	(.189)
creative		.315**		.192
		(.128)		(.124)
older siblings	.013	.019	.184	.185
	(.144)	(.145)	(.148)	(.146)
constant	.563	-.567	1.538*	.818
	(.741)	(.867)	(.887)	(1.012)
obs	92	92	92	92
pseudo R^2	.0786	.1329		

Table 1.7: Regression coefficients (robust standard errors between brackets) (* = significant at 10%; ** = significant at 5%; *** = significant at 1%); we use maximum likelihood estimation for our probit regressions (with the rationality indicator as dependent variable) and quasi-maximum likelihood estimation for our fractional response regressions (with the violation index as dependent variable)

If we first regard the control variables, we find that they all have the expected sign: being in the kindergarten has a negative effect on (the degree of) rationality, while having older siblings has a positive impact. However, the only significant effect is the one for age (i.e. the kindergarten dummy).

Let us then turn to our more interesting results on the different intellectual characteristics. First, if we consider our (main) specification 1, we find that mathematical ability has a positive and significant impact on rationality. In other words, if children outperform their peers in mathematics, then there is a higher probability that they take rational decisions. Interestingly, this result is robust for both the rationality indicator and the violation index as dependent variables. This is in line with Burks et al. (2009), who found that cognitive skills significantly impact on economic preferences and choices, favouring the individuals with better cognitive skills. Next, and perhaps somewhat surprisingly, we find that language has a strongly and significantly negative impact on being rational. Again, this finding holds for our two dependent variables. It appears that, on average, having better language ability goes together with less rational decision making.

An obvious question here is how we can square these results with our earlier findings (based on Table 1.6), where we concluded that neither mathematics nor language seemed to be related to rational behaviour. The explanation for this paradox is that, as we also mentioned above, Table 1.6 captures unconditional effects, whereas the regression results in Table 1.7 define conditional effects: our regressions effectively condition on the level of verbal skills when assessing the impact of non-verbal skills, and vice versa. Thus, to identify the positive effect of mathematical ability on rationality, it turns out to be important to control for the fact that language ability has an opposite effect on rationality. Because mathematical and language skills are positively correlated (i.e. children often perform well (or badly) in both mathematics and language simultaneously), one risks to miss the significant effects when conducting an unconditional analysis. In our opinion, this may also explain the result of Harbaugh et al. (2001) that we mentioned above. It seems that these authors found no effect of mathematical ability on the observed rationality of children's behaviour because they did not correct for differences in language skills.

To conclude, we consider the regression results for our specification 2, which includes creativity as an additional explanatory variable. Importantly, we observe that our significant effects for mathematical and language skills remain (roughly) the same as in our regressions that use specification 1. Next, the (positive) effect of creativity is significant when explaining the (binary) rationality indicator. However, the effect is not confirmed in the regression of the violation index. Remember that the effects of mathematical skills and language skills are stable across all four specifications. This leads us to conclude that language skills (as indicator of verbal intelligence) and mathematical skills (as indicator of non-verbal intelligence) are the more important drivers of rational consumption behaviour by children. Interestingly, this coincides with the theory of Gardner (1983), who argues that there are distinct types of intelligence, and that linguistic intelligence and logical-mathematical intelligence are very important in our society.

Finally, note that the regression results in Table 1.7 should be interpreted with care. First of all, the estimates from non-linear models (probit, glm) may suffer from omitted variable bias even when the omitted variable is not correlated with the included variables (see e.g. Yatchew and Griliches (1985)). Second, the coefficients are no longer significant when correcting for multiple-testing using the Holm-Bonferroni method[23]. On the other hand, measurement errors in the explanatory variables may cause the regression parameters to be biased.

1.6 Conclusion

Focusing on the consumption behaviour of children, we investigated the relationship between taking rational consumption decisions and being smart. To do so, we designed an experiment involving unsophisticated discrete consumption choices (with only three commodities). In addition, we developed a revealed preference test that is specifically designed for analysing the resulting choice data. Using teacher based assessments, we also obtained information on the children's personal characteristics, which we related to the children observed degree of rationality.

[23]Note however that the Holm-Bonferroni method is rather conservative.

The analysis of our experimental data obtained two main conclusions. A first conclusion is that, on average, children's consumption behaviour appeared not fully rational. In this respect, however, we also observed a clear age effect: children tend to behave more rational when growing older. More generally, we found quite some heterogeneity in rationality across children.

Our second important conclusion then pertains to relating this heterogeneity to specific child characteristics. Our particular focus here was on the effect of intellectual ability. Specifically, we considered indicators of verbal (language) and non-verbal (mathematical skills) intelligence. As a robust finding, we obtained that mathematical ability positively impacts on children's rationality, whereas language skills have a significantly negative effect. More generally, we take our results to mean that it is particularly important to acknowledge the multidimensional nature of 'being smart' when relating rationality to intelligence.

We believe that an interesting avenue for follow-up research consists of explaining our results in terms of the characteristics of the decision processes that underlie the observed consumption behaviour. For example, such an analysis may clarify our (perhaps somewhat surprising) result on the negative effect of language skills on the rationality of children's consumption decisions. Referring to our discussion in the Introduction, such an investigation can contribute further to a better protection of those children who appear to be particularly vulnerable consumers.

1.A Questions posed to teachers

For each of the dimensions taken up in Table 1.2 we asked the following questions to the teachers (translated from Dutch):

1. 'How would you position **[first name]** **[surname]** in terms of **[mathematical, language or creativity]** skills, when comparing **[him]**/**[her]** to all other pupils of similar age whom you have taught? Excellent (top 2%), Very good (top 10%), Good (top

25%), Average (top 50%), Less than average (bottom 50%), Bad (bottom 25%), Very bad (bottom 10%), Terrible (bottom 2%). (Indicate one answer, e.g. if you find that the pupil is good, but that he/she does not belong to the top 10%, indicate top 25%.)'

2. 'How would you rate the [**mathematical, language or creativity**] skills of [**first name**] [**surname**], if you compare [**him**]/[**her**] to average peers? (indicate one possibility)? 1 (very weak) to 10 (very strong)'

1.B Details on our experiment

1.B.1 Prices and choice sets

Tables 1.8 and 1.9 present the prices regimes that we used to construct our choice sets and the 9 discrete choice sets themselves.

	Prices	
1 unit of grapes	1 unit of mandarins	1 unit of letter biscuits
8	4	1
8	3	2
9	3	1
1	8	4
2	8	3
1	9	3
4	1	8
3	2	8
3	1	9

Table 1.8: The price regimes defining the choice sets

1.B.2 Experimental design

The children were welcomed one at a time in a separate room. Before starting the experiment, each child was allowed to taste the grapes, mandarins and letter biscuits. They were instructed that these products were similar to the products they could choose in a next step.

It was stated multiple times that each product was from the same brand and had the same quality.

We explained that they had to choose 9 successive times but that they would only receive one (randomly selected; see below) consumption bundle afterwards. Subsequently, each child was presented with the first of nine choice sets. He or she could choose one out of seven plates. Each plate displayed a consumption bundle consisting of a combination of grapes, mandarins and letter biscuits.

After a plate had been chosen by the child, we presented the next choice set, while again stating that his or her second choice was as important as the first one, and that choices were independent of each other. This process was repeated nine times. At the end of the experiment, the children were invited to draw a card with a number from one to nine. They received the corresponding consumption bundle.

Quantities		
Grapes	Mandarins	Biscuits
Choice 1		
1.5	0	0
0	3	0
0	0	12
0.5	1	4
0.75	0	6
0.75	1.5	0
0	1.5	6
Choice 2		
1.5	0	0
0	4	0
0	0	6
0.5	1.33	2
0.75	0	3
0.75	2	0
0	2	3
Choice 3		
1.33	0	0
0	4	0
0	0	12
0.44	1.33	4
0.66	0	6
0.66	2	0
0	2	6
Choice 4		
0	1.5	0
0	0	3
12	0	0
4	0.5	1
6	0.75	0
0	0.75	1.5
6	0	1.5
Choice 5		
0	1.5	0
0	0	4
6	0	0
2	0.5	1.33
3	0.75	0
0	0.75	2
3	0	2
Choice 6		
0	1.33	0
0	0	4
12	0	0
4	0.44	1.33
6	0.66	0
0	0.66	2
6	0	2
Choice 7		
0	0	1.5
3	0	0
0	12	0
1	4	0.5
0	6	0.75
1.5	0	0.75
1.5	6	0
Choice 8		
0	0	1.5
4	0	0
0	6	0
1.33	2	0.5
0	3	0.75
2	0	0.75
2	3	0
Choice 9		
0	0	1.33
4	0	0
0	12	0
1.33	4	0.44
0	6	0.66
2	0	0.66
2	6	0

Table 1.9: The 9 discrete choice sets

Chapter 2

Revealed preference theory for finite choice sets[1]

2.1 Introduction

Finite choice sets Revealed preference theory was initially developed to deal with situations where choices are made from linear budget sets. Although the theory has been extended to deal with nonlinear budgets (See, for instance, Yatchew (1985), Matzkin (1991), Forges and Minelli (2009), Cherchye et al. (2014)), none of these papers looks at the situation where choice sets are finite. However, in many settings choice sets are inherently finite.[2]

A first and obvious case where finite choice sets occur naturally is when the goods under consideration can only be bought in discrete amounts. In such cases, the choice set can be represented as the intersection of a linear budget set and the space $\mathbb{Z}_+^n$. This particular setting is studied in a recent paper by Polisson and Quah (2013). They show that for such choice sets, the standard revealed preference condition (i.e. GARP) characterises the data

[1] This chapter is based on joint work with Thomas Demuynck (Maastricht University). I refer to the working paper version of Cosaert and Demuynck (2013) and the article version of Cosaert and Demuynck (2014b).

[2] Recently Forges and Iehlé (2013) discussed the revealed preference conditions when the available data only consist of a so called 'essential experiment' given by observed consumption bundles and a feasibility matrix. In such setting the experimental observer only knows to which extent a bundle that has been chosen at some date is also available at another date.

sets which are rationalisable by a utility function which is separable in an unobserved good which can be consumed in continuous quantities. Independently of our work, Forges and Iehlé (2014) generalise the results of Polisson and Quah (2013) by characterising the necessary and sufficient revealed preference conditions for rationalisability when the choice set is given by the intersection of a budget set and some discrete set. In this way, the authors identify a counterpart of the standard GARP condition for situations where (all) goods are indivisible, which they call the Discrete Axiom of Revealed Preference. Their analysis differs from ours in two ways. First of all, our focus is on more general discrete choice settings, i.e. we do not necessarily require that the choice sets are constructed as the intersection of a discrete set and a linear budget set. Second, we have a different focus in the sense that we put strong emphasis on the additional properties that are imposed on the utility functions, like monotonicity and concavity. Nevertheless, there are also some clear connections between our results and the findings from Forges and Iehlé (2014). For example, if we restrict our setting to their framework, then the Discrete Axiom of Revealed Preference can be shown to be equivalent to our Weakly Monotone Axiom of Revealed Preference.[3]

A second setting where discrete choice sets in combination with revealed preferences are pertinent is for experimental data. Indeed, revealed preference theory is remarkably well suited to analyse the rationality of subjects in experimental settings. Its main advantage lies in the fact that experiments can be specifically designed to allow for very powerful tests (e.g. by letting prices vary and keeping budgets constant across different choice problems). The usual procedure for such revealed preference experiments is to let the subjects of the experiment solve a number of different exercises. For each exercise, the subject is endowed with a budget (expressed as a number of tokens) and is informed on a vector of prices. Next, the subjects are instructed to allocate their budget over a set of goods subject to the budget constraint defined by the income and the prices. See, for example, Cox (1997), Andreoni and Miller (2002), Février and Visser (2004), Fisman, Kariv, and Markovits

[3] See Section 2.3 for a formal definition.

(2007), Huck and Rasul (2008), Bruyneel, Cherchye, and De Rock (2012b) and Cherchye, Demuynck, and De Rock (2013a) for similar experimental designs. However, this experimental design may pose two potential problems. First of all, it requires that the subjects understand the concepts of money, prices and income and that they are able to compute the total expenditure and compare it with the total available budget. This requirement is not always satisfied, especially when the subjects under consideration are children (as in Chapter 1). Second, revealed preference theory usually requires that the entire available budget must be exhausted (i.e. the total expenditure should be equal to the available budget). In settings where there are only two goods, this requirement can be met by representing choice problems graphically as a 2–dimensional budget line. This setting can even be extended to include choices under uncertainty (Choi, Fisman, Gale, and Kariv (2007)). Then, budget exhaustion can easily be imposed by restricting the choices to lie on the budget line. However, if there are more than two goods, graphical illustrations are no longer feasible (or much more difficult to represent). As such, the requirement that the entire budget must be exhausted might require a lot of fine–tuning on the part of the subject (See Mattei (2000) for a clear illustration of this problem). In these settings, subjects are usually given calculators (or other computing devices) to check whether there are any tokens left to spend. Nevertheless, this fine–tuning might still impose a considerable burden on the subjects. In fact, this burden might become large enough such that it actually interferes with the optimality of the choices. Indeed, if there are many goods, then the opportunity cost (i.e. additional time) that is needed to fine–tune the choices may become quite large. This, in turn, might lead to situations where the subjects choose to lower the time spent on fine–tuning at the expense of choosing a less–optimal bundle.

An elegant solution to the two aforementioned problems is to design the experiment in such a way that the subjects choose from a finite set of distinct bundles. Restricting the choice sets to be finite makes the choices of the respondents much easier: they only have to pick one bundle from a finite collection of feasible options. The first experimental study that uses this option is by Harbaugh, Krause, and Berry (2001) who investigated the rationality of choice behaviour by children and young adults. Their experimental design has been replicated by several others (see for instance Burghart, Glimcher, and Lazzaro (2012) and List and Millimet (2008)). Another study that uses experimental data was discussed in Chapter 1. We will also use these data sets in our empirical illustration.

A third relevant case where choice sets are finite is when choices are made by picking a single item from a finite set of distinct alternatives (e.g. the choice between different cars from a catalogue). In these settings, it is useful to think of the different alternatives as representing different bundles of characteristics (e.g. the price, the top speed, the fuel efficiency, etc.). Usually, these kind of discrete choice models are analysed by econometric methods which are based on limited dependent variables models (see for instance Train (2009) for a thorough overview or Berry, Levinsohn, and Pakes (1995) for a seminal contribution). Given this, our results can actually be seen as a first step towards an analysis of such discrete choice characteristics models by nonparametric revealed preference techniques.[4] Related to this, we also like to point to the large choice theoretic (and behavioural) literature that models the choice behaviour over arbitrary (discrete) sets of alternatives, which need not necessarily be representable as bundles of goods or characteristics. The main rationalisability concept in this setting is due to Richter (1966), who provided a choice theoretic analogue to the 'consumption based revealed preference literature' founded by Samuelson (1938, 1948) and Houthakker (1950). By developing a 'consumption based' revealed preference theory for finite choice sets, we are in a certain sense building a bridge between these two largely separate literatures, thereby opening the door for empirical applications of various other choice theoretic models.

Contributions Our paper has several contributions. First of all, from a theoretical perspective, we derive a number of revealed preference conditions that can be applied to settings

[4] The revealed preference conditions of the characteristics model in a continuous choice space were analysed by Blow, Browning, and Crawford (2008).

where choices are made from finite sets of distinct consumption bundles. Towards this end, we distinguish between four cases: rationalisability by a weakly monotone utility function, rationalisabililty by a weakly monotone and concave utility function, rationalisability by a strongly monotone utility function and rationalisability by a strongly monotone and concave utility function. Here, rationalisability by a weakly monotone utility function is the weakest concept while rationalisability by a strongly monotone and concave utility function is the strongest. For each of the four rationalisability concepts we obtain a different set of revealed preference conditions. Interestingly, these different revealed preference conditions do not coincide. As such, it is for example possible to find a data set that is rationalisable by a weakly monotone utility function but not by a concave and strongly monotone utility function. This result is interesting because such distinctions can not be made when choices are obtained from linear budget sets. Indeed, a well known result in revealed preference theory (Afriat's theorem) tells us that in such cases, rationalisability by a weakly monotone utility function is equivalent to rationalisability by a strongly monotone and concave utility function. This means that when choice sets are linear, all four rationalisability concepts coincide (see Section 2.2 for more details).

Second, we provide a number of conditions on the finite choice sets for which it is allowed to neglect the fact that choices are made from finite choice sets. In other words, we present a collection of assumptions such that the standard revealed preference condition (i.e. GARP) is still valid for consistency with utility maximising behaviour. These conditions can be used to design experimental settings for which the results can still be analysed using the standard revealed preference conditions. For example, we show that the experimental design of Harbaugh, Krause, and Berry (2001) satisfies the conditions such that GARP characterises the data sets that are consistent with utility maximisation by a strongly monotone (and concave) utility function. However, we also show that it is not possible to strengthen this to utility maximisation with a weakly monotone (and concave) utility function. In other words, it possible that the data set violates GARP, but the behaviour was nevertheless generated by some weakly monotone utility function (e.g. a Leontief utility function).

Finally, we illustrate the relevance of our results by analysing two experimental data sets that contain choices made by children and young adults. The first data set is from the previously mentioned experiment of Harbaugh, Krause, and Berry (2001). The second is from Bruyneel, Cherchye, Cosaert, De Rock, and Dewitte (2012a) (Chapter 1 in this dissertation). For both data sets, we find that imposing weak monotonicity instead of strong monotonicity (and concavity) on the utility functions improves empirical fit in terms of higher predictive success, a measure which was recently introduced by Beatty and Crawford (2011).

In Section 2.2, we give a brief summary of the most important results in revealed preference theory. This discussion will be useful to position our results within the revealed preference literature. Section 2.3 contains the main theoretical results of this paper. It develops the revealed preference conditions for the finite choice set framework. In Section 2.4, we present some conditions for which the usual revealed preference conditions coincide with our revealed preference conditions. Section 2.5 contains an empirical illustration of our results. Finally, Section 2.6 concludes. All proofs are in Appendix 2.A.

2.2 Revealed preference theory for linear budget sets

In this section, we present the basic revealed preference theory. This will be useful for comparison with the results that will be established in Sections 2.3 and 2.4.

To start, consider a finite collection of sets $\{B_t\}_{t \in T}$, where T is a finite set of observations, $T = \{1, 2, \ldots, |T|\}$. In this section, we assume that the choice sets take on the form of a linear budget set,

$$B_t = \{\mathbf{q} \in \mathbb{R}^n_+ | \mathbf{p}_t \mathbf{q} \leq m_t\},$$

In words, the choice set B_t contains all bundles $\mathbf{q} \in \mathbb{R}^n_+$ that can be bought with a certain income $m_t > 0$ at prices $\mathbf{p}_t \in \mathbb{R}^n_{++}$. In the next sections, we will consider the setting where each budget set B_t consists of a finite number of distinct bundles.

A data set $S = \{B_t, \mathbf{q}_t\}_{t\in T}$ then consists of a finite number of budget sets and for each budget set B_t a bundle $\mathbf{q}_t$ from this set, i.e. $\mathbf{q}_t \in B_t$. The idea is that $\mathbf{q}_t$ is the bundle which is chosen from the budget set B_t. Usually, it is assumed that $\mathbf{q}_t$ lies on the boundary of B_t, i.e. $\mathbf{p}_t\mathbf{q}_t = m_t$. In most experimental settings, this condition is additionally imposed. Given this, data sets are also frequently written as $\{\mathbf{p}_t, \mathbf{q}_t\}_{t\in T}$ where the underlying budget set is implicitly defined by:

$$B_t = \{\mathbf{q} \in \mathbb{R}^n_+ | \mathbf{p}_t\mathbf{q} \leq \mathbf{p}_t\mathbf{q}_t\}.$$

The following defines the standard rationality concept in revealed preference theory.

Definition 2.1 (Rationalisability)**.** A data set $S = \{B_t, \mathbf{q}_t\}_{t\in T}$ is **rationalisable** by the utility function $u : \mathbb{R}^n_+ \to \mathbb{R}$ if for all $t \in T$,

$$\mathbf{q}_t \in \arg\max_{\mathbf{q}\in B_t} u(\mathbf{q}).$$

In words, a data set S is rationalisable by the utility function u if for each observation $t \in T$, the chosen bundle $\mathbf{q}_t$ maximises the utility function $u(.)$ over the budget set B_t.

For two vectors $\mathbf{q}$ and $\mathbf{q}'$, we write $\mathbf{q} \geq \mathbf{q}'$ if every element of the vector $\mathbf{q}$ is as least as large as the corresponding element of the vector $\mathbf{q}'$. We denote $\mathbf{q} > \mathbf{q}'$ if $\mathbf{q} \geq \mathbf{q}'$ and $\mathbf{q} \neq \mathbf{q}'$ and we write $\mathbf{q} \gg \mathbf{q}'$ if every element of $\mathbf{q}$ is strictly larger than the corresponding element of $\mathbf{q}'$. A utility function $u(.) : \mathbb{R}^n_+ \to \mathbb{R}$ is **weakly monotone** if $\mathbf{q} \geq \mathbf{q}'$ implies $u(\mathbf{q}) \geq u(\mathbf{q}')$ and $\mathbf{q} \gg \mathbf{q}'$ implies $u(\mathbf{q}) > u(\mathbf{q}')$. A utility function $u : \mathbb{R}^n_+ \to \mathbb{R}$ is **strongly monotone** if $\mathbf{q} \geq \mathbf{q}'$ implies $u(\mathbf{q}) \geq u(\mathbf{q}')$ and $\mathbf{q} > \mathbf{q}'$ implies $u(\mathbf{q}) > u(\mathbf{q}')$. The utility function $u : \mathbb{R}^n_+ \to \mathbb{R}$ is **locally non–satiated** if for every open neighbourhood N of $\mathbf{q}$ there is a bundle $\mathbf{q}' \in N \cap \mathbb{R}^n_+$ such that $u(\mathbf{q}') > u(\mathbf{q})$. Strong monotonicity is stronger than weak monotonicity (as it rules out situations like Leontief utility functions) which, in turn, is stronger than local non–satiation. The utility function $u : \mathbb{R}^n_+ \to \mathbb{R}$ is **concave** if for all $\mathbf{q}$ and $\mathbf{q}' \in \mathbb{R}^n_+$ and all $\alpha \in [0, 1]$, $u(\alpha \mathbf{q} + (1 - \alpha)\mathbf{q}') \geq \alpha u(\mathbf{q}) + (1 - \alpha)u(\mathbf{q}')$.

Given these properties on the utility function, it is possible to define different rationalisability concepts, e.g. rationalisability by a strongly monotone and concave utility function, or rationalisability by a weakly monotone utility function. As will be demonstrated in Theorem 2.2 below, if budget sets are linear, then all these rationalisability concepts coincide. However, as we will demonstrate in the following sections, this equivalence breaks down when choice sets are finite.

Theorem 2.2 (Afriat's theorem)**.** *Consider a data set $S = \{B_t, \mathbf{q}_t\}_{t \in T}$ where each set B_t ($t \in T$) is a linear budget set. Then the following conditions are equivalent:*

1. *The data set S is rationalisable by a locally non–satiated utility function.*

2. *The data set S satisfies GARP,*

3. *For all $t \in T$, there exist numbers ϕ_t and $\lambda_t > 0$ such that for all $t, v \in T$,*

$$\phi_t - \phi_v \leq \lambda_v \mathbf{p}_v (\mathbf{q}_t - \mathbf{q}_v).$$

4. *The data set S is rationalisable by a concave, (continuous) and strongly monotone utility function.*

The above theorem shows that rationalisability by a locally non–satiated utility function is equivalent to GARP. Next, the equivalence between the second and fourth condition states that GARP is also equivalent to rationalisability by a concave, strongly monotone and continuous utility function. The equivalence between the first and fourth condition actually shows that it is impossible to reject concavity and strong monotonicity of the utility function without rejecting utility maximisation by a locally non–satiated (and, hence, weakly monotone) utility function. In other words, if budget sets are linear, then all the different rationalisability concepts which are nested between the properties of 'local non–satiation' and 'strict monotonicity and concavity' coincide. In the next section, we will show that this property no longer holds if the choice sets are finite. In other words, the equivalence between the first and fourth condition turns out to be a consequence from the fact that choice sets take the shape of linear budget sets. The linear inequalities in the third condition are the so called Afriat inequalities. These have a nice interpretation when the underlying rationalisation is concave. Indeed, if, for example, $S = \{B_t, \mathbf{q}_t\}_{t \in T}$ is rationalisable by a concave and strongly monotone utility function $u(.)$ and if we take the simplifying assumption that $u(.)$ is differentiable, then concavity of $u(.)$ implies that for all t and v,

$$u(\mathbf{q}_t) - u(\mathbf{q}_v) \leq \nabla_{\mathbf{q}} u(\mathbf{q}_v)(\mathbf{q}_t - \mathbf{q}_v). \tag{2.1}$$

Here $\nabla_{\mathbf{q}} u(\mathbf{q}_v)$ is the gradient of $u(.)$ at the bundle $\mathbf{q}_v$. Strict monotonicity requires that $\nabla_{\mathbf{q}} u(\mathbf{q}_v) \gg \mathbf{0}$. Next, the first order conditions for the utility maximisation problem imply that,

$$\nabla_{\mathbf{q}} u(\mathbf{q}_v) \leq \lambda_v \mathbf{p}_v, \tag{2.2}$$

where λ_v is the strictly positive Lagrange multiplier corresponding to the budget constraint. Notice that we can replace the inequality sign with an equality sign for strictly positive quantities. If we substitute equality (2.2) into inequality (2.1) and if we set $\phi_t = u(\mathbf{q}_t)$ and $\phi_v = u(\mathbf{q}_v)$, we effectively obtain the Afriat inequalities. The Afriat inequalities form a set of linear inequalities. As such, they provide a second set of conditions by which it can be verified whether a data set is rationalisable.

2.3 Revealed preference theory for finite choice sets

In the previous section we considered the case where each choice set B_t takes on the form of a linear budget set. However, as explained in the introduction, in many contexts, individuals choose by picking one out of a finite number of distinct consumption bundles. To model this setting, we assume from now on that each choice set B_t consists of a finite number of distinct bundles $B_t = \{\mathbf{b}_t^1, \ldots, \mathbf{b}_t^{K_t}\} \in \prod_{i=1}^{K_t} \mathbb{R}_+^n$. Here $K_t = |B_t|$, which may depend on the observation $t \in T$.

As in the previous section, we denote by $\mathbf{q}_t$ the observed choice from the set B_t, i.e. $\mathbf{q}_t \in \{\mathbf{b}_t^1, \ldots, \mathbf{b}_t^{K_t}\}$ and we denote a data set S as $\{B_t, \mathbf{q}_t\}_{t \in T}$. The concept of rationalisability in this setting is identical to the definition of the previous section: the data set $S = \{B_t, \mathbf{q}_t\}_{t \in T}$ is rationalisable by the utility function $u(.)$ if for all $t \in T$,

$$\mathbf{q}_t \in \arg\max_{\mathbf{q} \in B_t} u(\mathbf{q}).$$

In the linear budget case, consumers maximise utility over all bundles in the linear budget set, which is defined by a price vector and a budget. This is reflected in the construction of the revealed preference relations. If, for instance, $\mathbf{q}_v$ is on the boundary of the budget set associated with $\mathbf{q}_t$ (i.e. $\mathbf{p}_t\mathbf{q}_v = \mathbf{p}_t\mathbf{q}_t$), we note that $u(\mathbf{q}_t) \geq u(\mathbf{q}_v)$. If $\mathbf{q}_v$ is in the interior of this budget set (i.e. $\mathbf{p}_t\mathbf{q}_v < \mathbf{p}_t\mathbf{q}_t$), this condition can be strengthened to $u(\mathbf{q}_t) > u(\mathbf{q}_v)$ because, by local non–satiation, there must be a bundle $\widetilde{\mathbf{q}}_t$ in a small neighbourhood of $\mathbf{q}_v$ such that $u(\mathbf{q}_t) \geq u(\widetilde{\mathbf{q}}_t) > u(\mathbf{q}_v)$. In the case of finite choice sets, on the other hand, rationalisability by a locally non–satiated utility function is no longer useful. The reason is that choice sets do not contain open subsets. For this reason, we will restrict ourselves to the two most natural strengthenings of local non–satiation, namely weak and strong monotonicity. These assumptions allow us to construct similar revealed preference relations in a discrete choice setting. After all, the standard GARP is sufficient (but not necessary) for rationalisability of choices from finite choice sets.

As indicated in the previous section, when choices are made from finite sets, the different rationalisability concepts no longer coincide. We will make a distinction between four different notions of rationalisability: (i) rationalisability by a weakly monotone utility function, (ii) rationalisability by a strongly monotone utility function, (iii) rationalisability by a weakly monotone and concave utility function and (iv) rationalisability by a strongly monotone and concave utility function. The gap between rationalisability by a weakly monotone utility function on the one hand and rationalisability by a strongly monotone and concave utility function on the other hand is of particular interest. The other rationalisability concepts are nested between these characterisations, imposing either strong monotonicity (ii) or concavity (iii), hence providing us with a tool to analyse the different assumptions on preferences in more detail.

In principle, it is possible to obtain for all four rationalisability concepts, revealed preference restrictions both in terms of GARP-like conditions and in terms of 'Afriat-like' inequalities. However, we will mainly focus on GARP-like conditions for the first two rationalisability concepts (the cases without concavity) and Afriat-like conditions for the rationalisability concepts with concavity. The reason for doing this is twofold. First, from a theoretical viewpoint, it turns out that the conditions that we present are the most intuitive for the rationalisability concept under consideration. The other revealed preference conditions are much more difficult to interpret. Second, from an empirical viewpoint, it turns out that the omitted revealed preference conditions are computationally much more difficult to verify and therefore less useful in practice.

For the remainder of this section, we first present rationalisability by a weakly (strongly) monotone utility function. Rationalisability by a weakly monotone utility function is the least restrictive rationalisability concept. Next, we focus on rationalisability by a strongly (weakly) monotone and concave utility function. Rationalisability by a strongly monotone and concave utility function imposes the strongest restrictions, i.e. the data must be rationalisable by a well-behaved utility function.

Weakly (strongly) monotone rationalisability Let us start with rationalisability where the underlying utility function is not required to be concave. For a particular budget set B_t, consider its comprehensive hull,

$$C(B_t) = \{\mathbf{q} \in \mathbb{R}^n_+ | \exists j \leq K_t : \mathbf{q} \leq \mathbf{b}_t^j\}$$

The set $C(B_t)$ contains all bundles which are dominated by some bundle in B_t. Of course, if a consumer has a monotone utility function, then the chosen bundle $\mathbf{q}_t$ must be at least as good as every bundle in $C(B_t)$. Given this, the main idea is to reformulate the basic revealed preference relations (used in the definition of GARP) using the sets $C(B_t)$ as the new budget sets.

Definition 2.3. A data set $S = \{B_t, \mathbf{q}_t\}_{t\in T}$ satisfies the **Weakly (resp. Strongly) Monotone Axiom of Revealed Preference** if there exist binary relations R and R_0 such that for all $t, v \in T$,

1. If there exists a bundle $\mathbf{q}_v \in C(B_t)$ then $\mathbf{q}_t R_0 \mathbf{q}_v$.
2. If there exist observations a, b and s such that $\mathbf{q}_t R_0 \mathbf{q}_a$, $\mathbf{q}_a R_0 \mathbf{q}_b$, ..., $\mathbf{q}_s R_0 \mathbf{q}_v$, then $\mathbf{q}_t R \mathbf{q}_v$.
3. If $\mathbf{q}_t R \mathbf{q}_v$ then for all $\mathbf{b}_v^k \in B_v$ it is not the case that $\mathbf{b}_v^k \gg \mathbf{q}_t$ (resp. $\mathbf{b}_v^k > \mathbf{q}_t$).

We say that $\mathbf{q}_t$ is directly revealed preferred to $\mathbf{q}_v$, $\mathbf{q}_t R_0 \mathbf{q}_v$ if $\mathbf{q}_v \in C(B_t)$ (or equivalently, if there exists a bundle $\mathbf{b}_t^j \in B_t$ such that $\mathbf{b}_t^j \geq \mathbf{q}_v$). As usual the indirect revealed preference relation R is defined as the transitive closure of the direct revealed preference relation. Then, we say that the data set $S = \{B_t, \mathbf{q}_t\}_{t\in T}$ satisfies the Weakly Monotone Axiom of Revealed Preference (WMARP) if $\mathbf{q}_t R \mathbf{q}_v$ implies that for all $\mathbf{b}_v^k \in B_v$ it is not the case that $\mathbf{b}_v^k \gg \mathbf{q}_t$. On the other hand, the data set $S = \{B_t, \mathbf{q}_t\}_{t\in T}$ is said to satisfy the Strongly Monotone Axiom of Revealed Preference (SMARP) if $\mathbf{q}_t R \mathbf{q}_v$ implies that for all $\mathbf{b}_v^k \in B_v$ it is not the case that $\mathbf{b}_v^k > \mathbf{q}_t$.

Given these definitions of WMARP and SMARP, we can state our first result. All the proofs of the theorems can be found in Appendix 2.A.

Theorem 2.4. *Consider a data set $S = \{B_t, \mathbf{q}_t\}_{t \in T}$. Then the following conditions are equivalent:*

- *The data set S is rationalisable by a weakly (resp. strongly) monotone (and continuous) utility function.*
- *The data set S satisfies the* ***Weakly (resp. Strongly) Monotone Axiom of Revealed Preference WMARP (resp. SMARP)****.*

Both conditions WMARP and SMARP can easily be verified by using a simple adaptation of the algorithm presented by Varian (1982) which is the standard method to verify GARP.

The definitions of SMARP and WMARP were obtained by adapting the definition of GARP to the budget sets $C(B_t)$. However, it is also possible to give conditions which are more similar to the Afriat type inequalities (See also the proof of Theorem 2.4 in Appendix 2.A for more details). In particular, let $\mathbf{e}_i$ be the n dimensional unit vector which has its i–th coordinate equal to one and which has all other coordinates equal to zero. Next let

$$a_{v,t} = \min_{k \leq K_v} \left(\max_i \mathbf{e}_i(\mathbf{q}_t - \mathbf{b}_v^k) \right)$$

It can easily be shown that $a_{v,t} \leq 0$ if and only if $\mathbf{q}_t \in C(B_v)$. It turns out that the data set $\{B_t, \mathbf{q}_t\}_{t \in T}$ satisfies WMARP if and only if there exist numbers ϕ_t and $\lambda_t > 0$ such that for all t, v,

$$\phi_t - \phi_v \leq \lambda_v a_{v,t}.$$

These inequalities are similar to the standard Afriat inequalities where the expression $\mathbf{p}_v(\mathbf{q}_t - \mathbf{q}_v)$ is replaced by $a_{v,t}$. Similar to the term $\mathbf{p}_v(\mathbf{q}_t - \mathbf{q}_v)$, we have that $a_{v,t}$ is less than or equal to zero if and only if $\mathbf{q}_t$ is in the budget set ($C(B_v)$) of observation v.

Strongly (weakly) monotone and concave rationalisability In this paragraph, we focus on our rationalisability concepts that involve concave utility functions. Concavity implies that there exists a supporting hyperplane associated with the indifference curve through each chosen bundle. When budget sets are linear, on the one hand, the lower half-space defined by this supporting hyperplane contains all 'revealed worse' bundles (i.e., all bundles in the budget set). When choice sets are finite, on the other hand, the slope of the indifference curve through each point in the choice set is unknown. The construction of one hyperplane through the chosen bundle is no longer sufficient to guarantee that the chosen bundle is 'revealed preferred' over all bundles in the comprehensive hull of the choice set. We therefore consider the convex indifference curve through each alternative $\mathbf{b}_t^k$ from the choice set B_t. As the utility function is concave, there should be a supporting hyperplane associated with this indifference curve through $\mathbf{b}_t^k$. Let us indicate its slope by $\mathfrak{p}_t^k (= \nabla_{\mathbf{q}} u(\mathbf{b}_t^k))$. We have that $\mathfrak{p}_t^k \gg 0$ if the utility function is strongly monotone (and $\mathfrak{p}_t^k > 0$ if the utility function is weakly monotone). Each bundle $\mathbf{b}_t^k$ is then the utility maximising alternative over all bundles $\mathbf{q}$ for which $\mathfrak{p}_t^k \mathbf{b}_t^k \geq \mathfrak{p}_t^k \mathbf{q}$. It follows that the data set $\{\mathfrak{p}_t^j, \mathbf{b}_t^j\}_{t \in T, j \leq K_t}$ should satisfy GARP.

In terms of Afriat inequalities, this means that there should exist numbers ϕ_t^j and $\mathfrak{p}_t^j$ such that for all $t, v \in T$ and all $k \leq K_t, j \leq K_v$:[5]

$$\phi_t^k - \phi_v^j \leq \mathfrak{p}_v^j (\mathbf{b}_t^k - \mathbf{b}_v^j). \tag{2.3}$$

We still need to link these inequalities to the actual observed choices. Recall that the number ϕ_t^k can be interpreted as the utility level corresponding to the bundle $\mathbf{b}_t^k$. Then, given that $\mathbf{q}_t = \mathbf{b}_t^k$ was chosen while another bundle $\mathbf{b}_t^j$ was also feasible at $t \in T$, it must be that $\phi_t^k = u(\mathbf{b}_t^k) = u(\mathbf{q}_t) \geq u(\mathbf{b}_t^j) = \phi_t^j$. Hence, we also require that for all observations $t \in T$ and all $j \leq K_t$, if $\mathbf{q}_t = \mathbf{b}_t^k$, then $\phi_t^k \geq \phi_t^j$. This implies that the actually chosen bundle $\mathbf{q}_t$

[5] We can omit the variables λ_v as the vectors $\mathfrak{p}_v^j$ are only defined up to scale.

is *not* below the supporting hyperplanes associated with alternatives $\mathbf{b}_t^j$.

Definition 2.5. A data set $S = \{B_t, \mathbf{q}_t\}_{t \in T}$ satisfies the **Strongly (resp. Weakly) Monotone and Concave Axiom of Revealed Preference** if for all $t \in T$ and $\mathbf{b}_t^k \in B_t$ there exist numbers ϕ_t^k and vectors $\mathfrak{p}_t^k \gg 0$ (resp. $\mathfrak{p}_t^k > 0$) such that for all $t, v \in T$:

$$\phi_t^k - \phi_v^j \leq \mathfrak{p}_v^j(\mathbf{b}_t^k - \mathbf{b}_v^j) \text{ and,}$$

$$\text{if } \mathbf{q}_t = \mathbf{b}_t^k \text{ then, } \phi_t^k \geq \phi_t^j \text{ for all } j \leq K_t.$$

Consistency with WMCARP (SMCARP) turns out to be a necessary and sufficient condition for rationalisability by a concave utility function.

Theorem 2.6. *Consider a data set $S = \{B_t, \mathbf{q}_t\}_{t \in T}$. Then the following conditions are equivalent:*

- *The data set S is rationalisable by a strongly (resp. weakly) monotone (continuous) and concave utility function.*
- *The data set S satisfies the* ***Strongly (resp. Weakly) Monotone and Concave Axiom of Revealed Preference SMCARP (resp. WMCARP)****.*

Although the first set of inequalities can be interpreted as Afriat inequalities, they deviate from the usual Afriat inequalities as given in Theorem 2.2 in three important ways. First, for WMCARP and SMCARP, we are faced with a total of $(\sum_t K_t)^2$ linear inequalities rather than T^2 inequalities for the usual Afriat inequalities. The reason is that for WMCARP (SMCARP) we have a price vector for every available bundle and not just for every chosen bundle. Second, for the Afriat inequalities in Theorem 2.6, the price vectors $\mathfrak{p}_t^k$ are variables. This is due to the fact that we do not observe the slope of the indifference curve through the bundles $\mathbf{b}_t^k$. Finally, the conditions WMCARP and SMCARP contain an additional set of restrictions: the utility associated with a chosen bundle must be at least as high as the utility associated with other alternatives in the same choice set.

Both SMCARP and WMCARP are expressed as a set of linear inequalities.[6] As such, they can easily be verified using linear programming methods. However, the conditions WMCARP and SMCARP can also be stated in revealed preference terms. Indeed, given our discussion above, we have that the vectors $\mathfrak{p}_v^j$ actually correspond to the prices at which the bundle $\mathbf{b}_v^j$ would have been chosen (from a linear budget set). As such, we need that $\{\mathfrak{p}_v^j, \mathbf{b}_v^j\}_{j \leq K_v, v \in T}$ must satisfy GARP. To take into account the second set of restrictions, this GARP condition should be complemented with the condition that no bundle $\mathbf{b}_t^j \in B_t$ is strictly (indirectly) revealed preferred to the chosen bundle $\mathbf{q}_t$. However, as the vectors $\mathfrak{p}_v^j$ are not observable, verifying the GARP condition is much more difficult than simply verifying the above set of linear inequalities.[7] That is the main reason why we choose to state it in this form.

2.4 GARP for finite choice sets

In the previous section, we characterised the different rationalisability concepts that can be applied when choice sets are finite. In this section, we present a number of sufficient conditions on the finite budget sets $\{B_t\}_{t \in T}$ such that GARP is still a necessary and sufficient condition for rationalisability. There are several reasons why this might be interesting.

First, despite the fact that our results in the previous section may provide a benefit for researchers who wish to distinguish between the different rationalisability concepts, it might equally well generate undesirable richness for researchers who are simply interested in the question whether subjects are rational or not. In this perspective, the conditions on the budget sets could be used to design an experiment with finite choice sets for which the usual GARP condition can still be applied to analyse the observed choice behaviour. Second,

[6]The condition that $\mathfrak{p}_v^j > \mathbf{0}$ can be met by requiring that the sum of the elements in $\mathfrak{p}_v^j$ should be strictly positive.

[7]If prices are unobserved, GARP can be tested using integer programming methods, see, for example, Cherchye, De Rock, and Vermeulen (2011a) and Cherchye, Demuynck, De Rock, and De Witte (2013b) for applications of such integer programming approach.

the standard GARP test is very well known and very easy to implement (and it is readily available in different programming languages). A researcher may therefore prefer to stick to this readily available test to verify rationalisability. Third, the standard GARP test relies on the use of price vectors. This facilitates the interpretation of deviations from rationality in terms of monetary loss (see, for example, Echenique, Lee, and Shum (2011) and Choi, Kariv, Müller, and Silverman (2014)). Finally, from a theoretical perspective, it might also be interesting to look at the conditions (on the budget sets) such that the different revealed preference characterisations from the previous section coincide with the usual GARP condition. In this way, we obtain a set of conditions for which the discrete choice setting is empirically indistinguishable from the linear budget setting.

In order to apply GARP, we need to introduce price vectors. The most straightforward way to do this is to assume that all bundles in the finite budget set B_t lie on some common budget hyperplane. In such case, there should exist vectors $\mathbf{p}_t \in \mathbb{R}^n_{++}$ and numbers m_t such that for all $t \in T$ and $j \in K_t$, $\mathbf{p}_t \mathbf{b}_t^j = m_t$. In a more general framework, one could also allow for choice sets where one or several options are dominated by some other option. However, in such setting, dominated alternatives should never be chosen because they imply a violation of monotonicity. As such, we maintain the assumption that the above equality is satisfied for all bundles $\mathbf{b}_t^j \in B_t$.

We consider two further conditions that may be satisfied for such budget sets.

Assumption 2.7. For all $t, v \in T$ and all $\mathbf{b}_v^k \in B_v$:

1. if $m_t \geq \mathbf{p}_t \mathbf{b}_v^k$, then there exists a bundle $\mathbf{b}_t^j \in B_t$ such that, $\mathbf{b}_t^j \geq \mathbf{b}_v^k$.
2. if $m_t > \mathbf{p}_t \mathbf{b}_v^k$, then there exists a bundle $\mathbf{b}_t^j \in B_t$ such that, $\mathbf{b}_t^j > \mathbf{b}_v^k$.

Assumption 2.8. For all $t, v \in T$ and all $\mathbf{b}_v^k \in B_v$:

1. if $m_t \geq \mathbf{p}_t \mathbf{b}_v^k$, then there exists a bundle $\mathbf{b}_t^j \in B_t$ such that, $\mathbf{b}_t^j \geq \mathbf{b}_v^k$.
2. if $m_t > \mathbf{p}_t \mathbf{b}_v^k$, then there exists a bundle $\mathbf{b}_t^j \in B_t$ such that, $\mathbf{b}_t^j \gg \mathbf{b}_v^k$.

The first condition of Assumptions 2.7 and 2.8 requires that if b_v^k is in the 'linearised' budget set of observation t, then it should also be in the set $C(B_t)$. This condition guarantees that the revealed preference relation for the GARP condition coincides with the revealed preference relation for the WMARP and SMARP conditions. The second condition of both assumptions requires that if b_v^k is in the interior of the 'linearised' budget set of observation t, then it should be strictly dominated by some bundle in B_t.

The following shows that under Assumption 2.7, GARP and SMARP coincide and that under Assumption 2.8 GARP is equivalent to WMARP.

Theorem 2.9. *If Assumption 2.7 (resp. 2.8) is satisfied, then a data set S satisfies GARP if and only if it satisfies SMARP (resp. WMARP).*

Assumption 2.7 guarantees that SMARP implies GARP. However, (by Theorem 2.2) GARP implies rationalisability by a strongly monotone and concave utility function, hence it also implies SMCARP. This, in turn, is stronger than SMARP. As such, Assumption 2.7 actually shows that GARP, SMARP and SMCARP will be indistinguishable. Similarly, we see that under Assumption 2.8, WMARP implies GARP. This, in turn, implies consistency with SMARP, SMCARP and WMCARP and all three are stronger than WMARP. In other words, under Assumption 2.8, all revealed preference tests, WMARP, SMARP, WMCARP and SMCARP, coincide. Theorem 2.9 also shows that it is the linearity of the budget sets, rather than the cardinality, that is responsible for the different rationalisability concepts to coincide under GARP.[8]

We present some illustrations (Figures 2.1 to 2.3) of the relevance of Assumptions 2.7 and 2.8. Each of the examples contains two choice sets $B_1 = \{\mathbf{b}_1^1, \mathbf{b}_1^2\}$ and $B_2 = \{\mathbf{b}_2^1, \mathbf{b}_2^2\}$ with two alternatives per choice set. We assume that $\mathbf{b}_1^1$ is chosen from B_1 ($\mathbf{q}_1 = \mathbf{b}_1^1$) and $\mathbf{b}_2^2$ is chosen from B_2 ($\mathbf{q}_2 = \mathbf{b}_2^2$).

Then we consider three scenarios. In the first scenario, depicted in Figure 2.1, Assump-

[8] We thank an anonymous referee for pointing this out for us.

tion 2.8 holds. We already discussed that when Assumption 2.8 holds, it is no longer possible to distinguish between WMARP, GARP and all rationality tests nested between WMARP and GARP. In the left panel of Figure 2.1, the choices can be rationalised. The dashed line presents an indifference curve (from a strongly monotone and concave utility function) that rationalises the data. Obviously, the data is also consistent with all weaker rationality tests. In the right panel, the choices cannot be rationalised. It is not possible to find even a weakly monotone and non-concave utility function that rationalises the observations.

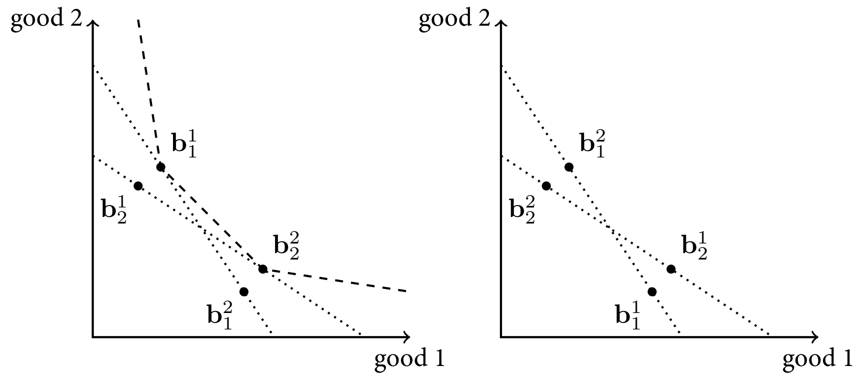


Figure 2.1: Illustration of choice sets that satisfy Assumption 2.8

In the second scenario, depicted in Figure 2.2, the choice sets satisfy Assumption 2.7 but violate Assumption 2.8. In this setting, one can distinguish between rationalisability with a weakly monotone and concave utility function (WMCARP) on the one hand and a weakly monotone and non-concave utility function (WMARP) on the other hand. While the choices in the right panel can be rationalised by a weakly monotone and concave utility function, the choices in the left panel can only be rationalised by a weakly monotone and non-concave utility function. In other words, the data in the left panel satisfy WMARP but not WMCARP.

Finally, Assumption 2.7 is violated in Figure 2.3. This makes that we can, in principle, make a distinction between rationalisability by a strongly monotone and concave utility

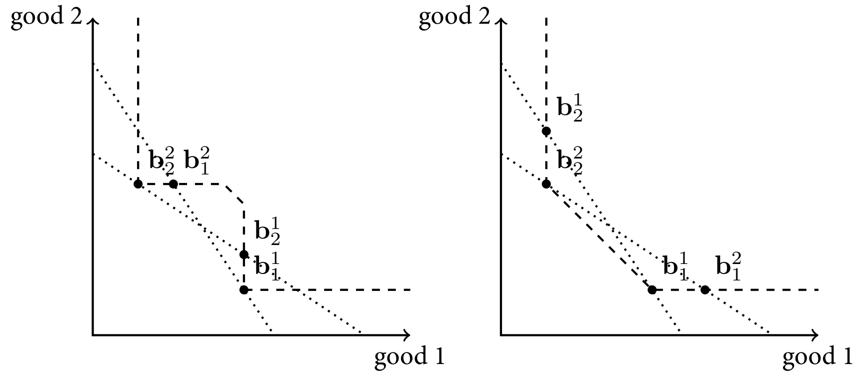


Figure 2.2: Illustration of choice sets that satisfy Assumption 2.7

function (SMCARP) on the one hand and a strongly monotone and non-concave utility function (SMARP) on the other hand. The left panel in Figure 2.3 depicts choices which satisfy SMARP but violate SMCARP. The choices in the right panel satisfy SMCARP.

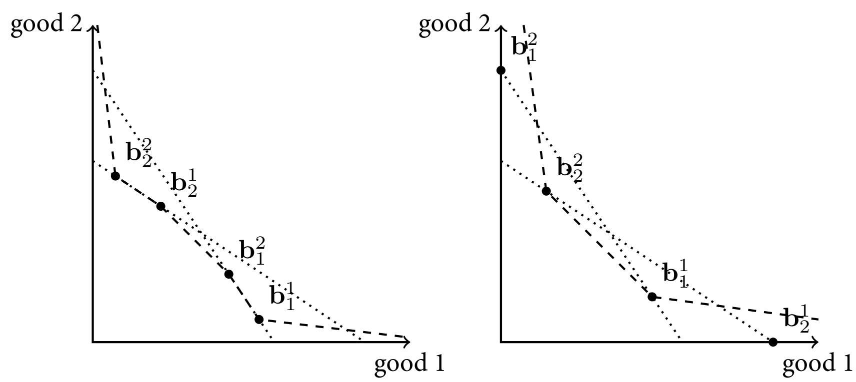


Figure 2.3: Illustration of choice sets that violate Assumption 2.7

These examples also indicate how to design experiments that allow maximum differentiation between subjects who satisfy WMARP, WMCARP, SMARP and SMCARP. Clearly, a setting in which choice sets violate Assumption 2.7 (Figure 2.3) has more potential for differentiation.

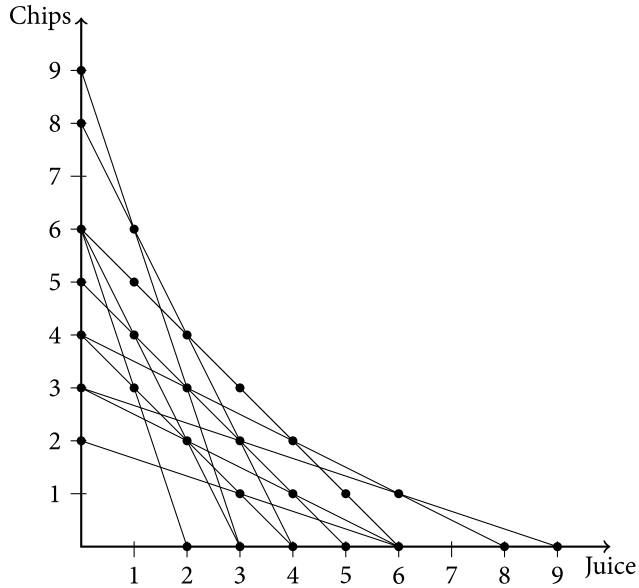


Figure 2.4: Experimental design of Harbaugh, Krause, and Berry (2001)

In order to see the practical relevance of the different assumptions, we take a closer look at the experimental design of Harbaugh, Krause, and Berry (2001) which will also be further analysed in the next section. Figure 2.4 presents the 11 different budget sets where each choice set consists of different bundles on the same budget line. It is easy to verify that these budget sets satisfy Assumption 2.7. As such, GARP will be equivalent to SMARP and SMCARP. This also means that it is impossible to distinguish between rationalisability by a strongly monotone utility function and rationalisability by a strongly monotone and concave utility function. On the other hand, the choice sets do not satisfy Assumption 2.8. As such, GARP will (in general) not be identical to WMARP and therefore, it might give us the opportunity to distinguish between rationalisability by a weakly monotone utility function and rationalisability by a weakly monotone and concave utility function. This feature will be demonstrated in the next section.

2.5 Illustration

We illustrate the usefulness of our results on the basis of two experimental data sets which use finite choice sets. The first is the data set from Harbaugh, Krause, and Berry (2001), which deals with choices made by children and young adults. The second is the data set from Bruyneel, Cherchye, Cosaert, De Rock, and Dewitte (2012a), which also deals with choices made by children. As mentioned in the introduction, letting children choose from linear budget sets is indeed problematic, because of difficulties that they may have in dealing with concepts like budgets and prices.
We first give a brief description of the two data sets (more information can be found in the respective papers). Next, we present the different measures by which we will compare the different tests: pass rate, power, predictive success and goodness–of–fit. Finally, we present and discuss our findings.

Brief dataset description The experiment from Harbaugh, Krause, and Berry (2001) (HKB from now on) contains information on 128 subjects (31 second grade students, 42 sixth grade students and 55 college undergraduates). Each subject had to choose from 11 different choice sets ($|T| = 11$). Each choice set contained different bundles of chips and juice (the choice sets are illustrated in Figure 2.4).

The second experiment, from Bruyneel, Cherchye, Cosaert, De Rock, and Dewitte (2012a) (BCCDD from now on), contains information about choice behaviour from 100 children (39 kindergarten respondents of about 5 years old, 31 third graders of about 8 years old and 30 sixth graders of about 11 years old). Each child was invited to solve nine successive choice problems ($|T| = 9$). Each choice set contained 7 distinct consumption bundles of three goods: grapes, mandarins and letter biscuits. The budget sets were chosen such that each bundle within the same budget lay, approximately, on the same hyperplane. The structure of these different hyperplanes can be found in Chapter 1.

Evidently, these experimental set-ups are highly artificial. However, both studies tried to increase the external validity of their results in a number of ways. First of all, during the experiments, the commodities were either physically presented on plates (BCCDD) or at least present in the room (HKB). Second, in BCCDD, children were invited to taste the grapes, mandarins and letter biscuits before the experiment started so the subjects had an ideal about their valuation of the different goods. In order for the revealed preference tests to be valid, we must also assume that the preferences of the subjects are stable across the different choice problems. This assumption of stable preferences may be problematic, particularly when it is imposed over longer periods. However, note that the time span of both experiments was rather short. Moreover, in order to exclude saturation effects, it was emphasised in both experiments that each choice is independent and equally important, and that subjects would only consume one of their choices at the end of the experiment (by randomly picking one of the chosen bundles as payoff). Although we acknowledge that the assumption of stable preferences is strong, we believe that preferences can be relatively stable in the setting under consideration. Nevertheless, it would be a bold claim to say that our results have strong external validity.

Although the data sets in HKB and BCCDD seem similar, there are also some important differences, which makes it interesting to compare the results between the two experiments. First of all, BCCDD let the subjects choose between different combinations of three commodities whereas HKB focus on combinations of two goods only. Given this, the choice problem in the BCCDD experiment is cognitively more challenging. Second, the respondents in BCCDD are on average younger than the respondents in HKB which may imply that they are less rational. Finally, the choice sets imposed by HKB satisfy Assumption 2.7 (but violate Assumption 2.8) while the choice sets imposed by BCCDD violate both assumptions. We therefore know that for HKB SMARP and SMCARP coincide with the standard revealed preference condition (GARP). However, this is not necessarily the case for the data from BCCDD. As such, by using both data sets, we also provide a clear empirical illustration of our findings in Section 2.4.

Pass rate, power, predictive success and goodness–of–fit In order to assess the empirical performance of the different revealed preference tests, we rely on three measures: the pass rate, the power and the predictive success. Next, we also look at the Houtman-Maks–index as a measure of goodness–of–fit.

Pass rate. The pass rate gives the percentage of all subjects that pass a certain revealed preference test. A higher pass rate implies that more subjects have made choices that can be rationalised. However, it is important to take into account the nestedness of the different tests. In particular, every data set that satisfies SMCARP will also satisfy SMARP, WMCARP and WMARP, every data set that satisfies SMARP will also satisfy WMARP and every data set that satisfies WMCARP will also satisfy WMARP. The reason for this is simply that every data set which is rationalisable by a certain utility function, will also be rationalisable by a utility function with weaker properties.

Power. Since the main aim of revealed preference theory is to provide an accurate description of real consumer behaviour, it is often favourable to look at more restrictive models. The strictness of a revealed preference test is usually measured by the 'power' of the revealed preference test. Basically, the power of a revealed preference test measures the probability that the *null* hypothesis of utility maximising behaviour is rejected when the *alternative* hypothesis of random behaviour holds. In the General Introduction, I have argued that there are several methods to 'simulate' random behaviour.

The first method is the 'bootstrap' procedure which was illustrated in the previous chapter. The bootstrap power index samples data sets from the empirical distribution of *observed* choices. The selected bundles are not exactly random bundles, because they correspond to an actually chosen bundle by some respondent. Nonetheless, the bundles of different respondents (who are characterised by different preferences) are randomly combined to construct the new data sets.

The second method is set out in Bronars (1987). For linear budget sets, the Bronars power measure is computed by constructing a high number of random data sets by drawing from a uniform distribution on the budget hyperplane. Sampling bundles from the uniform distribution is an implementation of an idea put forward by Becker (1962) that says that irrational behaviour involves random choices. As argued by Andreoni, Gillen, and Harbaugh (2011), the Bronars power measure might be attractive because the alternative hypothesis of random behaviour can be seen as a minimally informative prior. In our finite choice setting, we follow a modified version of the Bronars procedure. First, we generate 1000 random data sets. Each random data set is constructed by drawing, from each choice set B_t ($t \in T$), a bundle $\mathbf{q}_t$ at random (using a uniform distribution on $\{\mathbf{b}_t^1, \ldots, \mathbf{b}_t^{K_t}\}$). This gives us 1000 random data sets $\{B_t, \mathbf{q}_t\}_{t \in T}$. Finally, power is computed as the percentage of these random data sets that fail the revealed preference test under consideration.

Predictive success. Using a similar reasoning as for the pass rates, it is easy to see that for two nested models, the power index corresponding to the more restrictive model will be higher than the power for the more general model: if a random data set violates the weaker test, then it will also violate the more restrictive test. This implies, for example, that the power index for WMARP will be lowest among all revealed preference tests. In order to avoid these conflicting findings (high pass rates together with low power or low pass rates together with high power), we use a measure that combines both pass rate and power into a single metric: the measure of predictive success. The properties of this measure were originally proposed by Selten (1991) in Selten's Theorem. Beatty and Crawford (2011) applied the predictive success measure to a revealed preference setting and simplified the proof of Selten's Theorem. Predictive success is easily computed as the difference between the pass rate and one minus the power:

$$\text{predictive success} = \text{ pass rate } - [1 - \text{ power}].$$

The predictive success measure is increasing in both the pass rate and the power. As such, higher predictive success implies that a model is more successful at describing observed behaviour compared to random behaviour. The interpretation of the predictive success measure is quite intuitive. In the best case scenario, both pass rate and power are equal to one, which gives us a predictive success of one. In such case, all observed data sets pass the revealed preference test while all random data sets are rejected by the test. In this sense, the revealed preference test is perfectly able to distinguish between actual and random behaviour. In the worst case scenario, both pass rate and power are equal to zero, which gives us a predictive success of minus one. In this case, all observed data sets are rejected by the revealed preference test while all random data sets pass the test. In other words, the model explains random behaviour perfectly but not the actual behaviour. In intermediate cases, the measure of predictive success is found somewhere between minus one and plus one. A predictive success above zero points to a test which describes the observed data sets better than pure random behaviour. A negative predictive success indicates a setting where the revealed preference test under consideration explains better random behaviour than observed behaviour. A predictive success of zero implies that the test explains random behaviour as well as the observed behaviour (i.e. the test is unable to discriminate between random and observed behaviour).

Goodness–of–fit. The revealed preference tests (WMARP, SMARP, WMCARP and SMCARP) tell us whether or not a data set is consistent with utility maximising behaviour for various conditions on the underlying utility function. However, as convincingly argued by Varian (1990), in many cases, nearly optimising behaviour is just as good as optimising behaviour. As such, it would be useful to have some indication that says how close a given data set is to being rationalisable if it violates the revealed preference conditions. Usually, nearly optimising behaviour is measured by using a goodness–of–fit measure. The most popular goodness–of–fit measure in the revealed preference literature is without doubt Afriat (1973)'s Critical Cost Efficiency Index discussed in the General Introduction.

Given that we have no linear budget sets, it is not possible to apply the Critical Cost Efficiency Index in our setting. Moreover, the focus of this chapter is on the comparison of various characterisations. We therefore need a robust measure which has a similar interpretation across the different characterisations. An interesting measure in this respect is the Houtman-Maks index (HM–index) (Houtman and Maks (see 1985)). The HM–index gives the size of the largest subset of observations (i.e. the largest subset of T) which is still consistent with the revealed preference conditions under consideration. For example, the HM–measure for GARP looks at the largest subset of T, say A, such that $\{B_t, \mathbf{q}_t\}_{t \in A}$ still satisfies GARP. The HM–measure is difficult to compute. In particular, the problem is known to be NP–hard (Houtman and Maks, 1985; Dean and Martin, 2008). However, by reformulating the conditions in terms of a binary programming problem, one can effectively compute this measure for the finite choice set case.[9] Although binary programming methods are known to have exponential worst time complexity, they can be solved relatively fast for small to moderately sized problems.

The interpretation of the HM–measure is that subjects sometimes make mistakes when choosing their optimal bundle (i.e. for some observations $t \in T$, subjects fail to choose the best bundle). An alternative viewpoint could be that individuals always choose the optimal bundle, but that they sometimes fail to take into account all available options when making a choice. In this case, a more natural goodness–of–fit measure would be to look at the size of the largest subsets of options such that the observed choices are still consistent with the revealed preference test. A so called 'alternative–based' HM–index.[10] Formally, this index computes the largest number $n = \sum_t |B'_t|$ where $B'_t \subseteq B_t$ ($\forall t \in T$) and such that $\{B'_t, \mathbf{q}_t\}_{t \in T}$ satisfies the relevant revealed preference test.

[9]The detailed binary programming problems can be found in the working paper version of this paper (Cosaert and Demuynck, 2013).

[10]We thank an anonymous referee for proposing this alternative goodness-of-fit measure.

Harbaugh et al.	pass rate	Bronars		Bootstrap	
		power	pred succ	power	pred succ
GARP	0.547	0.995	0.542	0.987	0.534
	[0.457;0.635]		[0.452;0.630]		[0.444;0.622]
WMARP	0.828	0.898	0.726	0.812	0.640
	[0.751;0.889]		[0.649; 0.787]		[0.563; 0.701]
WMCARP	0.719	0.989	0.708	0.975	0.694
	[0.632;0.795]		[0.621;0.784]		[0.607;0.770]
Bruyneel et al.					
GARP	0.40	0.969	0.369	0.915	0.315
	[0.303;0.503]		[0.271;0.472]		[0.218;0.418]
WMARP	0.71	0.705	0.415	0.594	0.304
	[0.611;0.796]		[0.316;0.501]		[0.205;0.390]
WMCARP	0.55	0.879	0.429	0.794	0.344
	[0.447;0.650]		[0.326;0.529]		[0.241;0.444]
SMARP	0.43	0.928	0.358	0.868	0.298
	[0.331;0.533]		[0.259;0.461]		[0.199;0.401]
SMCARP	0.40	0.933	0.333	0.888	0.288
	[0.303;0.503]		[0.236;0.436]		[0.191;0.391]

Table 2.1: Pass rates, power and predictive success

Results The first column of Table 2.1 presents pass rates associated with the observed data and the corresponding 95% confidence intervals. Further, the table also reports the Bronars and bootstrap power and the predictive success for the different revealed preference tests. Given that the data sets from HKB satisfy Assumption 2.7, we have that GARP is equivalent to SMARP and SMCARP. As such, we only report the results for GARP.

In order to compare the different revealed preference tests, one might be tempted to look directly at the difference in pass rates between the revealed preference tests. However, as the tests are nested, this might lead to wrong conclusions. For example, the pass rate for WMARP will always be higher than the pass rate of all other tests, but this does not necessarily mean that WMARP is also the best model to explain the observed choices. A better option is to compare the predictive success of the different tests. Table 2.2 provides results on several likelihood ratio tests for the null hypothesis of equal predictive success

Bronars	H_0	H_1	χ^2	p-value
Harbaugh et al.				
	WMARP $=$ WMCARP	WMARP $\neq$ WMCARP:	0.494	0.482
	WMARP $=$ GARP	WMARP $\neq$ GARP:	34.655	0.000
	WMCARP $=$ GARP	WMCARP $\neq$ GARP:	108.915	0.000
Bruyneel et al.				
	WMARP $=$ SMARP	WMARP $\neq$ SMARP:	1.7753	0.182
	WMARP $=$ SMCARP	WMARP $\neq$ SMCARP:	3.5516	0.059
	WMARP $=$ WMCARP	WMARP $\neq$ WMCARP:	0.1394	0.709
	SMARP $=$ SMCARP	SMARP $\neq$ SMCARP:	5.8139	0.016
	WMCARP $=$ SMCARP	WMCARP $\neq$ SMCARP:	12.458	0.000
	WMCARP $=$ SMARP	WMCARP $\neq$ SMARP:	3.7606	0.052
Bootstrap	H_0	H_1	χ^2	p-value
Harbaugh et al.				
	WMARP $=$ WMCARP	WMARP $\neq$ WMCARP:	2.987	0.084
	WMARP $=$ GARP	WMARP $\neq$ GARP:	8.792	0.003
	WMCARP $=$ GARP	WMCARP $\neq$ GARP:	79.69	0.000
Bruyneel et al.				
	WMARP $=$ SMARP	WMARP $\neq$ SMARP:	0.018	0.893
	WMARP $=$ SMCARP	WMARP $\neq$ SMCARP:	0.122	0.727
	WMARP $=$ WMCARP	WMARP $\neq$ WMCARP:	1.056	0.304
	SMARP $=$ SMCARP	SMARP $\neq$ SMCARP:	0.443	0.506
	WMCARP $=$ SMCARP	WMCARP $\neq$ SMCARP:	3.174	0.075
	WMCARP $=$ SMARP	WMCARP $\neq$ SMARP:	1.566	0.211

Table 2.2: Likelihood ratio tests for equality of predictive success

(see Appendix 2.B for more details on the construction of this likelihood ratio test).

Let us first have a look at the results from the experiment of HKB. We see that the pass rates of both WMARP (82%) and WMCARP (71%) are substantially higher than the pass rates of the GARP (SMARP, SMCARP) test (54%). Furthermore, these higher pass rates are not entirely offset by a lower Bronars power. As such, we see that the highest predictive success is for WMARP (0.726) and WMCARP (0.708). To confirm that the predictive success of WMARP and WMCARP is indeed higher than the predictive success of GARP (SMARP, SMCARP) (0.542), we can look at the results of the likelihood ratio test in Table 2.2. The results clearly suggest that both WMARP (p-value < 0.001) and WMCARP (p-value < 0.001)

have a significantly higher predictive success than GARP (SMARP, SMCARP) (although we cannot reject the hypothesis that WMARP and WMCARP have equal predictive success). The power and predictive success values based on the bootstrap approach confirm these findings.

Let us now look at the results for the experiment of BCCDD. These data sets violate Assumption 2.7. We already argued that these cases are extremely interesting. Beside discriminating between rationalisability by a weakly and strongly monotone utility function, we are also able to discriminate between situations where the data set is rationalisable by a (strongly or weakly) monotone utility function on the one hand and situations where the data set is rationalisable by a (strongly or weakly) monotone and concave utility function on the other hand. Also notice that in this setting, GARP does not necessarily correspond to any 'nice' rationalisability concept (although it is still a sufficient condition for rationalisability by a strictly monotone and concave utility function). Nevertheless, Table 2.1 also gives results for GARP as it is frequently used in the literature. Similar to the experiment of HKB, we see that the pass rates for WMARP (71%) and WMCARP (55%) are higher than for SMARP (43%) and SMCARP (40%), which is the lowest. Notice that the 43% pass rate corresponds to the (overall) pass rate found in Chapter 1. This is no surprise because the rationality test in Chapter 1 verified consistency with a strongly monotone utility function, which corresponds to SMARP. Then, if we also take into account the Bronars power, we see that WMCARP has the highest predictive success (0.429) of all models and is closely followed by WMARP (0.415). These numbers are higher than the predictive success of the other two models (0.358 for SMARP and 0.333 for SMCARP). We then look at the likelihood ratio tests for the equality of the various predictive success values. It turns out that the predictive success associated with SMCARP is significantly lower than the predictive success associated with WMARP (p-value $= 0.059$), WMCARP (p-value < 0.001) and SMARP (p-value < 0.02). Furthermore, the test also suggests that the predictive success of WMCARP is higher than that of SMARP (p-value $= 0.052$). Following the bootstrap approach, the differences are less outspoken. Only the positive difference between the predictive success scores of WMCARP and SMCARP is (border line) statistically significant.

Previously, we argued that BCCDD present a more complex decision problem. This is confirmed by the pass rates which are lower for the BCCDD data set than for the HKB data. However, we see that for both data sets, WMCARP is among the best performing models (in terms of higher predictive success) while SMCARP has the lowest predictive success. This indicates that our finding is robust with respect to the complexity of the decision process. The use of both data sets also illustrates the empirical meaning of Assumption 2.7. Because Assumption 2.7 holds for the HKB data sets, pass rates, power estimates and predictive success values coincide for SMARP, SMCARP and GARP. In this sense it is no longer possible to distinguish between strong monotonicity with and without concavity. This also implies that a (linear) GARP test can still be applied to the HKB data, providing a strongly monotone rationalisation of the data. Only in the HKB setting, the (linear) GARP corresponds to a meaningful rationalisability concept.

Data set		by observation HM		by alternative HM	
		WMARP	SMCARP	WMARP	SMCARP
Harbaugh et al.	mean	10.758	10.266	49.617	48.938
	median	11	11	50	50
	std dev	0.649	1.046	1.323	1.999
	min	6	5	38	35
	max	11	11	50	50
Bruyneel et al.	mean	8.57	7.86	61.88	60.17
	median	9	8	63	62
	std dev	0.769	1.181	2.152	4.008
	min	6	5	54	45
	max	9	9	63	63

Table 2.3: HM-measure per observation versus HM-measure per alternative

Let us now have a look at the goodness–of–fit of the different tests in terms of the HM–measure. Table 2.3 compares the results obtained from the HM–measure by observation with the results obtained from the HM–measure by alternative. On the one hand, the results show that the by alternative measure is somewhat more refined. Observe for instance the lowest values of the usual HM–measure for the BCCDD data. To rationalise all data sets by a weakly monotone utility function, the HM–measure must be equal to (or lower than) 6. In other words, at least 3 observations must excluded. To rationalise all data sets by a strongly monotone and concave utility function, the HM–measure must be equal to (or lower than) 5. Only one additional observation must be excluded. However, when focusing on the 'by alternative' HM–measure, we see that 9 bundles must be removed for all data sets to satisfy WMARP, whereas up to 18 bundles must be removed for all data sets to be consistent with SMCARP. On the other hand, both measures are clearly related. The correlation between the two measures is above 89% for both data sets. Our general conclusions with respect to goodness–of–fit remain unaffected, regardless of the goodness–of–fit measure used. Figures 2.5 and 2.6 give the distribution of the 'by observation' HM–index for the different revealed preference tests for the two experiments. The black histogram gives the distribution of the HM–measure for the randomly generated data sets which were also used for computing the Bronars power measures. As such, it gives the distribution of the HM–measure for the hypothesis of random behaviour. The gray histogram, on the other hand, gives the distribution of the HM–measure for the actual data sets. As can be seen, the distributions of the real data sets have much more mass at higher values of the HM–index compared to the distribution generated by the random data sets (a Chi-square test rejects the equality of the HM-distributions for the actual and random data at all levels of significance). This confirms our earlier findings that the pass rates of the simulated data sets are consistently below the pass rates of observed data sets. This again suggests that the hypothesis of random behaviour is rejected for all models under consideration. As expected, the distribution of the HM–measure for the weakest test (WMARP) is most skewed to the right, but so is the distribution of the HM–measure when WMARP is applied to the randomly generated data sets.

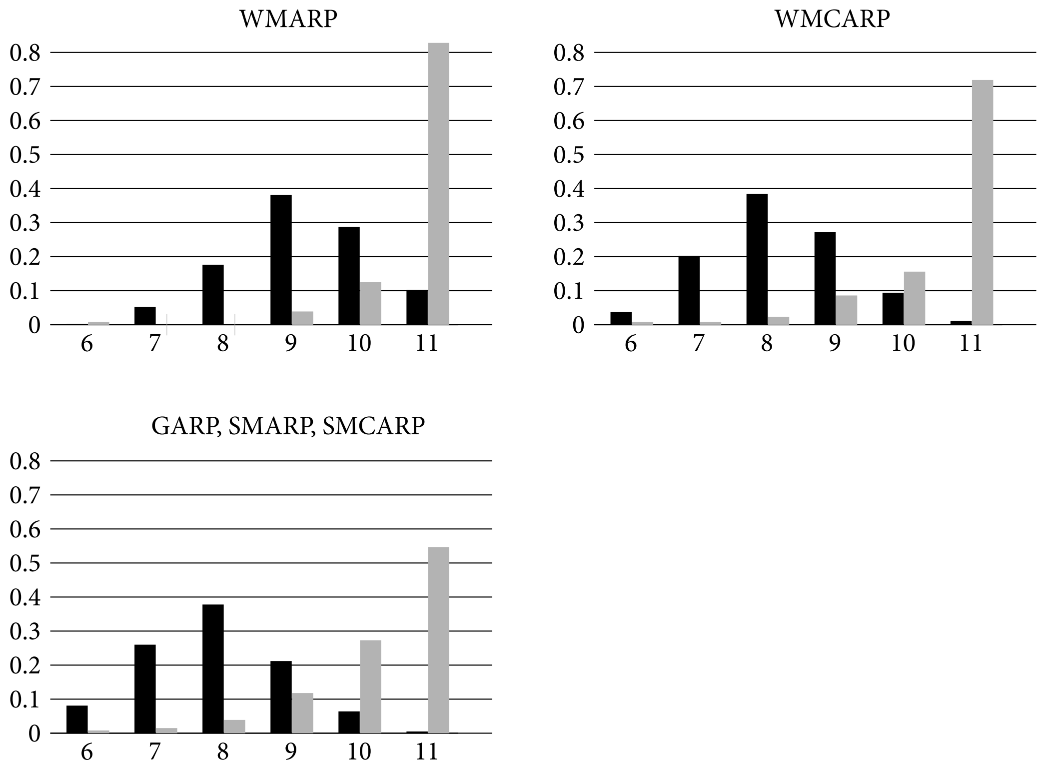


Figure 2.5: Distribution of HM–index for random and actual data for the data sets of Harbaugh et al.

What do we learn from all this? Our illustration has both a methodological and an empirical contribution. From a methodological point of view, we believe that it demonstrates the usefulness of our revealed preference characterisations to deal with choice models when choice sets are finite. It also shows how to use pass rates, power, predictive success and goodness–of–fit (HM–measure) to compare the empirical performance of the different revealed preference tests.

Next at the empirical level, we found that all four tests performed considerably better than a model which is based on pure random behaviour. Moreover, we have shown how various rationalisability concepts may lead to very different pass rates, power estimates and predictive success values. Both Harbaugh, Krause, and Berry (2001) and Bruyneel, Cher-

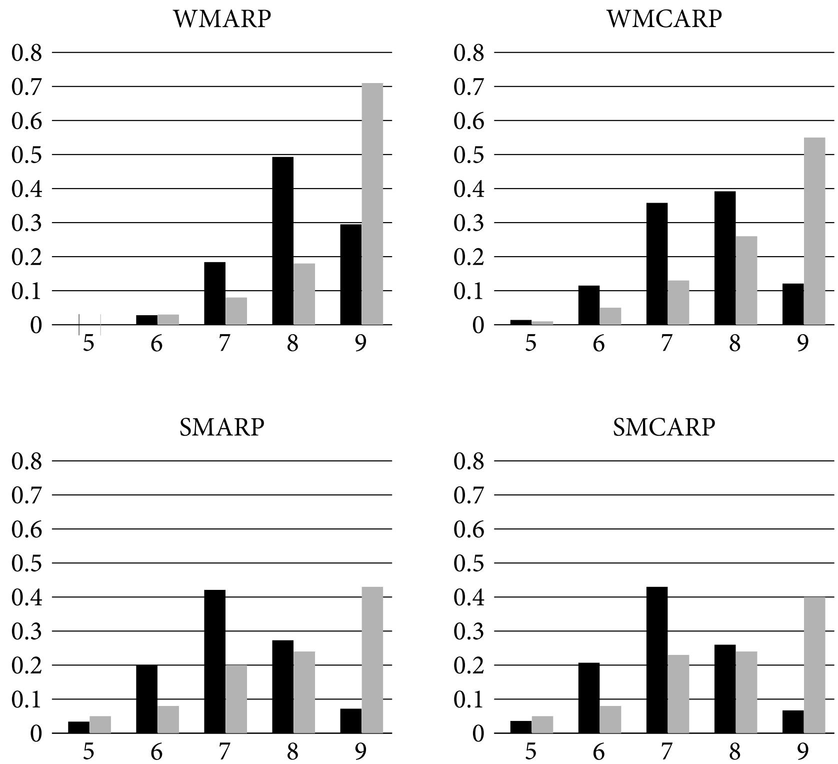


Figure 2.6: Distribution of HM–index for random and actual data for the data sets of Bruyneel et al.

chye, Cosaert, De Rock, and Dewitte (2012a) tested the rationality of children's consumption decisions based on a test equivalent to SMARP. HBK found that 55 per cent of the children behaved in a rational way and BCCDD found that only 43 per cent of the children were rational. We have shown how these pass rates can be improved up to 83 per cent and 71 per cent, respectively, by just dropping the strong monotonicity assumption (i.e. applying WMARP (or WMCARP) instead of SMARP)[11]. This shows that care should be taken with respect to the assumptions that are imposed on the underlying utility functions (i.e. strong vs weak monotonicity).

[11] The empirical relevance of the distinction between weak and strong monotonicity stems primarily from the large number of weakly dominated bundles across different choice sets.

2.6 Conclusion

We developed a revealed preference analysis for situations where choices are made from a finite collection of bundles. This setting occurs in various real life and experimental settings.

First of all, we have shown that when choices are made from finite choice sets, then different rationalisability concepts will have different revealed preference restrictions. This makes it possible to test for various conditions on the utility function, like strong monotonicity or concavity. Next, we have put forward a number of conditions for which our revealed preference conditions still coincide with the usual GARP condition. This result may be relevant for experimental researchers who do not wish to let their results depend on the specific conditions that are imposed on the utility function.

Finally, we applied our results using two experimental data sets that collect choices by children and young adults. We have shown that strong monotonicity may not be the best assumption to describe the observed choice behaviour.

We see several avenues for follow up research.

First of all, to focus our discussion, we have concentrated on testing rationalisability for basic conditions on the utility function, i.e. monotonicity and concavity. However, it is possible to obtain revealed preference conditions for even more stringent conditions on the utility function, like homotheticity or additivity (see Varian (1983) for such revealed preference conditions in the case of linear budget sets). A natural follow up research would be to derive the revealed preference conditions for such kind of utility functions when the choice sets are finite.

A second interesting subject for follow up research is the recovery or identification of the underlying preferences (or utility function) and to forecast behaviour in new choice situations (see Varian (1982) for recovery in the linear budget set setting). As for the setting considered in the paper, recovery could proceed using the revealed preference relation as obtained from the definitions of WMARP and SMARP or the 'utility' values of ϕ_t^j as obtained from the definitions of WMCARP and SMCARP.

A third avenue for follow up research is to investigate how our results can be useful for analysing (non-experimental) real life data. We have argued that finite choice sets occur naturally when the goods under consideration can only be bought in discrete amounts. Moreover, particular decisions (e.g. car purchase) can be represented as picking a single item from a set of alternatives with various characteristics. However, dealing with real life data involves both data measurement problems and unobserved preference heterogeneity. Extending our results to these settings will require methodological extensions which may build on recent work of Blundell, Browning, and Crawford (2003, 2008), Blundell, Kristensen, and Matzkin (2014), Hoderlein (2011), Blundell, Horowitz, and Parey (2013), and Hoderlein and Stoye (2014) who explicitly take into account individual heterogeneity for revealed preference tests with linear budget sets. Extending their insights to choice settings with finite budget sets would be very useful from a practical point of view as it would pave the way for convincing applications on the basis of real life data.

2.A Proofs

2.A.1 Proof of Theorem 2.4

Let us focus on weak monotonicity before presenting the strong monotonicity case. First, assume that $S = \{B_t, \mathbf{q}_t\}_{t \in T}$ is rationalisable by a weakly monotone utility function. We proceed by verifying that the data set satisfies WMARP.

Assume that $\mathbf{b}_t^k \geq \mathbf{q}_v$. Then obviously, by weak monotonicity, $u(\mathbf{b}_t^k) \geq u(\mathbf{q}_v)$. Next, as $\mathbf{q}_t$ was chosen from B_t, we also have that $u(\mathbf{q}_t) \geq u(\mathbf{b}_t^k)$. Therefore, $u(\mathbf{q}_t) \geq u(\mathbf{q}_v)$ and therefore we have that $\mathbf{q}_t R_0 \mathbf{q}_v$ implies $u(\mathbf{q}_t) \geq u(\mathbf{q}_v)$. By transitivity, we also have that $\mathbf{q}_t R \mathbf{q}_v$ implies $u(\mathbf{q}_t) \geq u(\mathbf{q}_v)$. Now, let $\mathbf{q}_t R \mathbf{q}_v$ (i.e. $u(\mathbf{q}_t) \geq u(\mathbf{q}_v)$) and assume, towards a contradiction, that $\mathbf{b}_v^k \gg \mathbf{q}_t$. Then as $\mathbf{q}_v$ was chosen from B_v, we have that

$u(\mathbf{q}_v) \geq u(\mathbf{b}_v^k)$. Next, from $\mathbf{b}_v^k \gg \mathbf{q}_t$ and weak monotonicity of the utility function, we have that $u(\mathbf{b}_v^k) > u(\mathbf{q}_t)$. A such, $u(\mathbf{q}_v) > u(\mathbf{q}_t)$. This contradicts with the assumption that $u(\mathbf{q}_t) \geq u(\mathbf{q}_v)$.

Now, assume that S satisfies WMARP. We need to show that it is also rationalisable by a weakly monotone utility function. Consider the n–dimensional unit vectors,

$$
\begin{aligned}
\mathbf{e}_1 &= (1, 0, \ldots, 0); \\
\mathbf{e}_2 &= (0, 1, \ldots, 0); \\
&\ldots \\
\mathbf{e}_n &= (0, 0, \ldots, 1);
\end{aligned}
$$

Next, define the functions $a_t : \mathbb{R}_+^n \to \mathbb{R}$ in the following way:

$$a_t(\mathbf{q}) = \min_{k \leq K_t} \left(\max_i \mathbf{e}_i(\mathbf{q} - \mathbf{b}_t^k) \right)$$

Note that the minimisation part of this function exploits feasibility: it minimises the maximum difference over the set of consumption bundles $\mathbf{b}_t^k$ that were feasible when $\mathbf{q}_t$ was chosen. This function satisfies the property that $a_t(\mathbf{q}) \leq 0$ if and only if there is a $k \leq K_t$ such that $\mathbf{q} \leq \mathbf{b}_t^k$ and $a_t(\mathbf{q}) < 0$ if and only if there is a $k \leq K_t$ such that $\mathbf{q} \ll \mathbf{b}_t^k$.

The function a_t is weakly monotone. Indeed if $\mathbf{q}' \geq (\gg)\mathbf{q}$, then

$$\max_i \mathbf{e}_i(\mathbf{q}' - \mathbf{b}_t^k) \geq (>) \max_i \mathbf{e}_i(\mathbf{q} - \mathbf{b}_t^k)$$

for all $i = 1, \ldots, n$ and therefore, $a_t(\mathbf{q}') \geq (>) a_t(\mathbf{q})$. Next, $a_t(\mathbf{q})$ is also continuous as it is given by the maximum of the minimum of continuous functions. We also have that for all $t \in T$, $a_t(\mathbf{q}_t) = 0$. Indeed, $a_t(\mathbf{q}_t) \leq 0$ because $\mathbf{q}_t \leq \mathbf{q}_t$. Now, if on the contrary $a_t(\mathbf{q}_t) < 0$ this would mean that there is a $k \leq K_t$ such that $\mathbf{q}_t \ll \mathbf{b}_t^k$, which would contradict WMARP.

Let $a_{t,v} = a_t(\mathbf{q}_v)$. We use the following definition of Cyclical Consistency.

Definition 2.10 (CC)**.** Consider a set of numbers $S = \{a_{t,v}\}_{t,v\in T}$. The set S is said to be cyclically consistent (CC) if there exists a binary relation W such that:

1. if $a_{t,v} \leq 0$, then tWv,

2. if tWv and vWw, then tWw,

3. if tWv then it is not the case that $a_{v,t} < 0$.

Lemma 2.11. The data set $S = \{B_t, \mathbf{q}_t\}_{t\in T}$ satisfies WMARP if and only if $\{a_{t,v}\}_{t,v\in T}$ satisfies CC.

Proof. Assume that S satisfies WMARP and let R be the indirect revealed preference relation. Assume that W is the relation as in the definition of CC. Let us first show that all three conditions of CC are satisfied if we take tWv if and only if $\mathbf{q}_t R \mathbf{q}_v$.

For the first, let $a_{t,v} \leq 0$. This means that $a_t(\mathbf{q}_v) \leq 0$ or equivalently $\mathbf{q}_v \leq \mathbf{b}_t^k$ for some $k \leq K_t$. However, this implies that $\mathbf{q}_t R_0 \mathbf{q}_v$ and, therefore, tWv. Hence, Condition 1 in CC is satisfied.

The second condition in the definition of CC follows immediately from the transitivity of the indirect revealed preference relation.

For the third condition let tWv which implies $\mathbf{q}_t R \mathbf{q}_v$. This implies that for no $k \leq K_v$, $b_v^k \gg \mathbf{q}_t$. Assume on the contrary that $a_{v,t} < 0$. This implies that there is a $k \leq K_v$ such that $\mathbf{e}_i(\mathbf{q}_t - \mathbf{b}_v^k) < 0$ for all $i = 1, \ldots, n$. However, this implies that $\mathbf{q}_t \ll \mathbf{b}_v^k$ which contradicts with the requirement of WMARP.

The proof that CC implies WMARP can be shown along the same lines. $\square$

Now, by a theorem of Fostel, Scarf, and Todd (2004) we have that CC is equivalent to the existence of numbers ϕ_t such that,

$$\phi_t - \phi_v \leq \lambda_v a_{v,t}.$$

Consider the function

$$u(\mathbf{q}) = \min_t \phi_t + \lambda_t a_t(\mathbf{q})$$

This function is continuous (as it is the minimum of continuous functions), it is weakly monotone (because $a_t(.)$ is weakly monotone for all $t \in T$) and we have that for all $t \in T$, $u(\mathbf{q}_t) = \phi_t$. In order to see this, notice that $u(\mathbf{q}_t) \leq \phi_t + \lambda_t a_t(\mathbf{q}_t) = \phi_t$. Now, if on the contrary $u(\mathbf{q}_t) < \phi_t$, then there must be an observation $v \in T$ such that $\phi_v + \lambda_v a_{v,t} < \phi_t$, a contradiction.

Now, let us show that u rationalises the data set. Assume, towards a contradiction that $u(\mathbf{b}_t^k) > u(\mathbf{q}_t)$ for some $k \leq K_t$. Then

$$\begin{aligned}
u(\mathbf{b}_t^k) &= \min_v u_v + \lambda_v a_v(\mathbf{b}_t^k), \\
&\leq u_t + \lambda_t a_t(\mathbf{b}_t^k), \\
&\leq u_t.
\end{aligned}$$

The last inequality comes from the fact that $\mathbf{b}_t^k \leq \mathbf{b}_t^k$, hence, $a_t(\mathbf{b}_t^k) \leq 0$. Let us now turn to the strong monotonicity case. Assume that the data set $S = \{B_t, \mathbf{q}_t\}_{t \in T}$ is rationalisable by a strongly monotone utility function u. Let us show that S satisfies SMARP. Similar to the weak monotonicity case, we can show that $\mathbf{q}_t R \mathbf{q}_v$ implies $u(\mathbf{q}_t) \geq u(\mathbf{q}_v)$. Now, assume that $\mathbf{q}_t R \mathbf{q}_v$, which implies $u(\mathbf{q}_t) \geq u(\mathbf{q}_v)$. If on the contrary $\mathbf{b}_v^j > \mathbf{q}_t$ for some $j \leq K_v$, then by strong monotonicity, $u(\mathbf{b}_v^j) > u(\mathbf{q}_t)$ and therefore, $u(\mathbf{q}_v) \geq u(\mathbf{b}_v^j) > u(\mathbf{q}_t)$, which gives us a contradiction.

To see the reverse, let $S = \{B_t, \mathbf{q}_t\}_{t\in T}$ satisfy SMARP. Consider the vectors $\mathbf{e}_i$ such that

$$\mathbf{e}_1 = (1, \varepsilon, \dots, \varepsilon);$$
$$\mathbf{e}_2 = (\varepsilon, 1, \dots, \varepsilon);$$
$$\dots$$
$$\mathbf{e}_n = (\varepsilon, \varepsilon, \dots, 1);$$

Here ε is a small but positive number.

Define the function a_t such that:

$$a_t(\mathbf{q}) = \min_{k \le K_t} \left(\max_i \mathbf{e}_i(\mathbf{q} - \mathbf{b}_t^k) \right).$$

Now, it is easy to see that if there is a $k \le K_t$ such that $\mathbf{q} \le \mathbf{b}_t^k$, then $a_t(\mathbf{q}) \le 0$. Also if there is a $v \in T$ such that $\mathbf{q}_v \not\le \mathbf{b}_t^k$ for all $k \le K_t$, we can set ε small enough such that $a_t(\mathbf{q}_v) > 0$. In other words, we can make ε small enough such that for all $v \in T$, $a_t(\mathbf{q}_v) \le 0$ if and only if $\mathbf{q}_v \le \mathbf{b}_t^k$ for some $k \le K_t$. Also, notice that $a_t(\mathbf{q}_t) = 0$ as otherwise $\mathbf{q}_t < \mathbf{b}_t^k$ for some $k \le K_t$, which contradicts SMARP. Also, the function a_t is easily seen to be strongly monotone and continuous.

Lemma 2.12. The set $\{a_{t,v}\}_{t,v}$ is cyclically consistent if and only if $\{B_t, \mathbf{q}_t\}_{t\in T}$ satisfies SMARP.

Proof. Let $\{B_t, \mathbf{q}_t\}_{t\in T}$ satisfy SMARP and let R be the revealed preference relation. Define the relation W such that tWv if and only if $\mathbf{q}_t R \mathbf{q}_v$. Let us show that W satisfies the definition of CC. First, let $a_{t,v} \le 0$. This means that there is a $k \le K_t$ such that $\mathbf{q}_v \le \mathbf{b}_t^k$. However, this implies that $\mathbf{q}_t R \mathbf{q}_v$ and therefore tWv as was to be shown. The second condition follows from transitivity of the relation R. For the third condition, let tWv which implies $\mathbf{q}_t R \mathbf{q}_v$. Now, if on the contrary $a_{v,t} < 0$, we know that there is a $k \leq K_v$ such that $\mathbf{q}_t \leq \mathbf{b}_v^k$. Now, if $\mathbf{q}_t = \mathbf{b}_v^k$, we have that $a_v(\mathbf{q}_t) = 0$, which is a contradiction. As such, it follows that $\mathbf{q}_t < \mathbf{b}_v^k$. However, this contradicts with SMARP. □

The remaining part of the proof is similar to that of the weakly monotone case.

2.A.2 Proof of Theorem 2.6

We only prove the Theorem for the strongly monotone case. The proof for the weakly monotone case is very similar. The first part of the proof is established in the text. It is shown that the first condition implies the second. For the reverse, let us assume that the data set $S = \{B_t, \mathbf{q}_t\}_{t \in T}$ satisfies SMCARP. Next, define the function

$$u(\mathbf{q}) = \min_{t \in T, k \leq K_t} \phi_t^k + \mathfrak{p}_t^k(\mathbf{q} - \mathbf{b}_t^k).$$

This function is continuous, concave and strongly monotone. Let us show that it rationalises the data. First of all, we show that $u(\mathbf{b}_t^k) = \phi_t^k$. The inequality $u(\mathbf{b}_t^k) \leq \phi_t^k$ follows simply from the definition of u. Now, if on the contrary $u(\mathbf{b}_t^k) < \phi_t^k$, then there must exist an observation $v \in T$ and $j \leq K_v$ such that $\phi_v^j + \mathfrak{p}_v^j(\mathbf{b}_t^k - \mathbf{b}_v^j) < \phi_t^k$. This contradicts SMCARP.

Now in order to show that $u(.)$ rationalises the data set $S = \{B_t, \mathbf{q}_t\}_{t \in T}$ assume, towards a contradiction, that there is a $t \in T$ and $k \leq K_t$ such that $u(\mathbf{b}_t^k) > u(\mathbf{q}_t)$. Then if $\mathbf{q}_t = \mathbf{b}_t^j$ it follows that $\phi_t^k > \phi_t^j$. However, this contradicts with the second condition of SMCARP.

2.A.3 Proof of Theorem 2.9

We only prove the theorem for Assumption 2.7. The proof corresponding to Assumption 2.8 is very similar.

Assume that the data set $S = \{B_t, \mathbf{q}_t\}_{t \in T}$ satisfies Assumption 2.7. Let us show that if S satisfies GARP, then S also satisfies SMARP. Let R_0 and R represent the direct and (indirect) revealed preference relations as given in the definition of GARP. We show that these relations also satisfy all conditions in the definition of SMARP. First, if $\mathbf{b}_t^k \geq \mathbf{q}_v$ for some $t \in T$ and $k \leq K_t$, we have that $m_t = \mathbf{p}_t \mathbf{b}_t^k \geq \mathbf{p}_t \mathbf{q}_v$ and therefore $\mathbf{q}_t R_0 \mathbf{q}_v$. Now, for the closing condition, assume on the contrary that $\mathbf{q}_t R \mathbf{q}_v$ and $\mathbf{b}_v^j > \mathbf{q}_t$ for some $v \in T$ and $j \leq K_v$. But then, $\mathbf{p}_v \mathbf{q}_v = \mathbf{p}_v \mathbf{b}_v^j > \mathbf{p}_v \mathbf{q}_t$ which violates GARP, a contradiction. Conclude that S satisfies SMARP.

For the other implication, assume that S satisfies SMARP and let R_0 and R represent the direct and (indirect) revealed preference relations that satisfy the definition of SMARP. We show that these relations also satisfy the definition of GARP. For the first condition, assume that $\mathbf{p}_t \mathbf{q}_t \geq \mathbf{p}_t \mathbf{q}_v$. However, by Assumption 2.7, this implies that $\mathbf{b}_t^k \geq \mathbf{q}_v$ for some $k \leq K_t$. As such, $\mathbf{q}_t R_0 \mathbf{q}_v$ as was to be shown. For the closing condition, let $\mathbf{q}_t R \mathbf{q}_v$ and assume on the contrary that $\mathbf{p}_v \mathbf{q}_v > \mathbf{p}_v \mathbf{q}_t$. From Assumption 2.7 this implies that $\mathbf{b}_v^j > \mathbf{q}_t$ for some $j \leq K_v$. However, this contradicts SMARP. As such, GARP must be satisfied.

2.B Likelihood ratio test for equal predictive success

To test the null hypothesis of equal predictive success across different characterisations, we make use of a likelihood ratio test that imposes linear restrictions on the proportion of a multinomial variable.

We distinguish between cases where the two revealed preference tests are nested and when they are not nested.

Likelihood ratio test for nested cases Assume that we want to test the equality of the predictive success between two nested models, for example WMARP and SMARP. First we transform the data into a multinomial variable.

- x_1 is the number of subjects that satisfy SMARP.
- x_2 is the number of subjects that satisfy WMARP but not SMARP.
- x_3 is the number of subjects that satisfy neither SMARP nor WMARP.

Let p_i be the probabilities of the respective outcomes. If the predictive success of SMARP is equal to the one of WMARP, then $p_1 - (1 - power_{SMARP}) = (p_1 + p_2) - (1 - power_{WMARP})$. Rewriting this gives $p_2 = power_{SMARP} - power_{WMARP}$. Let $A = power_{SMARP} - power_{WMARP}$. Then the restricted loglikelihood estimator solves:

$$\max_{p_1, p_2, p_3} \sum_i x_i \ln(p_i) \text{ s.t. } \sum_i p_i = 1, p_2 = A.$$

This gives the following maximum likelihood estimates:

$$\begin{aligned}
\tilde{p}_1 &= \frac{x_1(1-A)}{x_1 + x_3}, \\
\tilde{p}_2 &= A, \\
\tilde{p}_3 &= 1 - A - \frac{x_1(1-A)}{x_1 + x_3} = \frac{x_3(1-A)}{x_1 + x_3}.
\end{aligned}$$

On the other hand, the maximum likelihood estimators for the unrestricted model are equal to $\hat{p}_i = \frac{x_i}{\sum_i x_i}$. Given this we have that the test statistic:

$$-2 \ln \left(\sum_i x_i \ln(\tilde{p}_i) \right) + 2 \ln \left(\sum_i x_i \ln(\hat{p}_i) \right)$$

is asymptotically Chi–squared distributed with one degree of freedom in case the null hypothesis ($p_2 = A$) holds.

Likelihood ratio test for nonnested cases Let us now consider nonnested tests that compare, for example, the predictive success between SMARP and WMCARP. We consider the following multinomial variable:

- x_1 is the number of subjects that satisfy both SMARP and WMCARP,
- x_2 is the number of subjects that satisfy SMARP but not WMCARP,
- x_3 is the number of subjects that satisfy WMCARP but not SMARP,
- x_4 is the number of subjects that satisfy neither SMARP nor WMCARP.

Again letting p_i be the real proportions of the respective probabilities, we have that the null hypothesis of equal predictive success is equal to the condition that $p_1 + p_2 - (1 - power_{SMARP}) = p_1 + p_3 - (1 - power_{WMCARP}))$. Let $A = power_{SMARP} - power_{WMCARP}$. Then the hypothesis is equal to the condition that $p_3 = A + p_2$. The restricted loglikelihood estimator solves:

$$\max_{p_1,p_2,p_3,p_4} \sum_i x_i \ln(p_i) \text{ s.t. } \sum_i p_i = 1, p_3 = A + p_2.$$

After some calculations, we find that $\tilde{p}_2$ should solve the following quadratic equation:

$$\tilde{p}_2^2(-2(x_1+x_2+x_3+x_4))+\tilde{p}_2((1-3A)x_2+(1-A)x_3-2A(x_4+x_1))+A(1-A)x_2 = 0.$$

This equation has one positive root. Then $\tilde{p}_3 = \tilde{p}_2 + A$, $\tilde{p}_1 = x_1(1 - 2\tilde{p}_2 - A)/(x_4 + x_1)$, $\tilde{p}_3 = \tilde{p}_2 + A$ and $\tilde{p}_4 = 1 - \tilde{p}_1 - \tilde{p}_2 - \tilde{p}_3$. As before, we can use these values to evaluate the likelihood ratio test.

Part III

Psychological realism in economic modelling: identification of preferences for others' consumption and preferences for value

In this part, I focus on the recovery of marginal willingness to pay for others' consumption, and marginal willingness to pay for value. This requires an extension of the standard revealed preference approach. After all, the standard revealed preference tests of consistency with the neo-classical utility maximisation hypothesis are based on a narrow definition of consumers' preferences. Consumers are assumed to be purely self-interested, and only concerned with the quantities consumed.

I incorporate psychological realism in the revealed preference approach by modifying the arguments of the underlying utility functions. In Chapter 3, I let others' consumption enter the utility function of economic agents. Positive preferences for others' consumption may be driven by altruism (Andreoni and Miller (2002), Fisman et al. (2007) and Cox et al. (2008)), inequality-aversion (Fehr and Schmidt (1999) and Bolton and Ockenfels (2000)), concerns for efficiency and the pay-offs of the least well off (Charness and Rabin (2002) and Engelmann and Strobel (2004)) and reciprocity (Charness and Rabin (2002)). In this chapter, I specifically focus on collective consumption decisions made by children. When children make joint consumption choices, violations of rationality may be explained by deviations from the assumption of purely self-interested preferences. For this reason, I depart from the purely egoistic model of collective consumption to quantify externalities in consumption.

In Chapter 4, I let the utility function of consumers depend on their vector of expenditures. This allows us to capture so called diamond effects (Ng (1987)) which occur when some commodities are valued specifically for their value. Letting preferences depend on market prices seems to contradict the basic assumption of preference stationarity which is required for revealed preference analysis. However, I will show that expenditures can be incorporated as arguments of the utility function, while the utility function itself remains homogeneous. I find that the alternative revealed preference conditions are still refutable, and that it is possible to capture the fraction of the consumer's marginal willingness to pay that stems from her preferences for the value associated with some good.

Both chapters fit in Rabin (2013)'s PEEM research approach. These Portable Extensions of Existing Models allow for comparisons between some original model (in our case: egoistic and unitary models of rationality) on the one hand and alternative specifications corresponding to different behavioural assumptions on the other hand. In practice, this implies the introduction of an additional set of parameters which capture both the original model and its alternative specifications for given values of the parameters. Moreover, the parameters in this part have a convenient interpretation as they capture preferences for others' consumption and preferences for value in (monetary) terms of willingness to pay. Hence, these methods provide useful tools for identification in empirical analysis.

It is worth noting, however, that the focus of Chapters 3 and 4 is somewhat different. In Chapter 3 I focus on the interpersonal heterogeneity in preferences for others' consumption. I identify one 'selfishness' parameter per individual/dyad. In Chapter 4 I focus on heterogeneity in preferences for value associated with different commodities. The commodities under consideration are very diverse. Therefore, I identify one 'diamondness' parameter per commodity.

Chapter 3

Measuring the willingness to pay for others' consumption[1]

3.1 Introduction

This study is motivated by Rabin (2013)'s 'PEEM' (Portable Extensions of Existing Models) research program, which aims at developing tractable refinements of existing economic models that integrate psychological insights. The program encourages the design of new models that encompass a basic, pre-existing model at one particular parameter value, while other values for the same parameter imply modifications of the basic model. Rabin recommends the modelling of social preferences as a prime PEEMish application area. The literature has produced a mass of experimental evidence that rejects the standard model of purely selfish behaviour. However, Rabin argues that the replacing models with social preferences typically fail to derive plausible economic implications beyond specific laboratory environments. This indicates a need for analytical tools to handle non-selfish preferences in more general settings.

In the current paper, we apply Rabin's PEEM program to a specific type of social pref-

[1] This chapter is based on joint work with Sabrina Bruyneel (KU Leuven), Laurens Cherchye (KU Leuven), Bram De Rock (ULB) and Siegfried Dewitte (KU Leuven). I refer to the working paper version of Bruyneel et al. (2014).

erences (other types of social preferences are discussed in the concluding section). We consider the role of positive externalities in group consumption behaviour. In the case of group consumption, positive externalities imply that individuals are not purely selfish, but willing to pay for others' consumption. We introduce a methodology that allows us to measure this revealed willingness-to-pay in monetary terms. In line with our above motivation, this methodology associates a parameter value of unity with the standard model of purely selfish consumers, but also includes a whole range of other models (with varying externalities) for lower parameter values.

We apply our methodology to analyse the consumption choices made by dyads (i.e. two-person groups) of children in a tailored experiment. As we discuss in detail in Subsection 3.3, there is quite some debate in the literature on how (non-)selfish behaviour corresponds to specific child characteristics (in particular age). In our application, we first investigate to what extent children's consumption decisions are effectively characterised by externalities. Subsequently, we examine how age, gender and friendship between dyad members relate to revealed non-selfishness, so adding useful empirical input to the existing debate. At a more general level, this application shows the practical usefulness of our method to analyse the presence and determinants of non-selfish consumer behaviour.

The remainder of this introductory section specifies our research question. We also introduce the basic framework of our measurement methodology, and motivate our empirical application.

Non-selfish preferences. Consumer preferences are characterised by externalities when individual utilities depend not only on the own material consumption but also on the others' consumption.[2] In the empirical literature there is plenty of evidence that economic agents often act non-selfishly. For example, in social dilemma games, experimenters find that sub-

[2]Importantly, preferences with externalities are more general than caring preferences (in line with Becker (1974, 1981)), where a consumer's individual utility also depends on the others' aggregate utilities. Such caring preferences are a special case of the type of preferences we consider here; see Chiappori (1992) for more discussion.

jects cooperate even in one-shot games, when the only rational choice under selfishness is to defect; in ultimatum games subjects offer a substantial amount of tokens to their counterparties; in dictator games the dictators often share a fraction of their budget. The literature has suggested many alternative explanations for these phenomena, including altruism (Andreoni and Miller (2002), Fisman et al. (2007) and Cox et al. (2008)), inequality-aversion (Fehr and Schmidt (1999) and Bolton and Ockenfels (2000)), concerns for efficiency and the pay-offs of the least well off (Charness and Rabin (2002) and Engelmann and Strobel (2004)) and reciprocity (Charness and Rabin (2002)).

In the current paper, our focus is not on explaining the nature of the externalities, but instead on measuring the degree of externalities in a general setting of group consumption. To do so, we assume a structural model of rational group behaviour, which allows for consumption externalities and enables us to quantify the monetary value of externalities as individuals' willingness-to-pay for others' consumption. In particular, we can check how large this willingness-to-pay needs to be in order to rationalise the observed group consumption decisions. This methodology has several useful applications. For example, it can be used to quantify the extent to which models with selfish consumers are 'wrong' and so may lead to biased conclusions. Also, as we will illustrate in our own application, it allows us to relate the degree of externalities to specific consumer characteristics, so identifying which type of consumers is generally more or less selfish.

Measuring externalities. We assume the cooperative model as our structural model of group consumption (with and without selfish preferences). This consumption model was originally proposed by Apps and Rees (1988) and Chiappori (1988, 1992) and is nowadays widely used for analysing multi-person consumption behaviour. The model is particularly well-suited for addressing our research question, because it defines rational group consumption as a Pareto efficient allocation over group members. Importantly, this is the sole assumption that is made regarding the intra-group decision process. This reinforces the relevance of the empirical findings, as it avoids bias through additional, more debatable assumptions. In our particular context, a convenient implication of the Pareto efficiency assumption is that it allows us to define personalised prices to quantify consumption externalities in monetary terms. Specifically, these personalised prices reveal the willingness-to-pay of each group member for the own and the others' consumption.

Technically, to identify these personalised prices we will make use of a revealed preference methodology.[3] This methodology has a number of particularly attractive features within the present context. Most notably, it is intrinsically nonparametric, which means that it does not require a prior parametric/functional specification of the individual preferences. This minimises the risk that our empirical measurement of preference externalities (and the conclusions that are drawn from it) is confounded by some non-verifiable (and, thus, possibly erroneous) structure that is imposed on the consumption decision process. Next, from a practical point of view, the methodology evaluates rationality of group behaviour through testable conditions that are easily verified on data sets with a limited number of consumption choices (like in our application). Attractively, this also makes that the methodology does not need pooling of consumption data associated with different groups of consumers. The rationality of each group can be evaluated separately, which implies that we can maximally account for inter-group heterogeneity. Thus, our use of revealed preference methods avoids functional misspecification and debatable homogeneity assumptions, which effectively obtains a very 'pure' empirical assessment.[4]

Given our particular research interest, we define a new 'selfishness parameter' that cap-

[3]See Cherchye et al. (2007, 2011a) for revealed preference methodology to assess consumption decisions in terms of the cooperative consumption model. These authors build on early contributions of Samuelson (1938), Afriat (1967), Diewert (1973) and Varian (1982), who focused on rational (i.e. utility maximising) individual behaviour. Sippel (1997) argues that revealed preference methods are particularly useful in combination with experimental data such as the ones used in our own application. See also Harbaugh et al. (2001) and Bruyneel et al. (2012a), who use revealed preference methods to assess the rationality of children's individual consumption decisions.

[4]Similar motivations underlie the studies of Andreoni and Miller (2002), Fisman et al. (2007) and Cox et al. (2008), who also use revealed preference methods to study the altruistic behaviour of individual consumers. The research questions of these authors are closely related to the one that we consider here. However, a main difference is that we focus on the measurement of externalities in the context of group consumption decisions, while these other authors consider the presence of altruism under individual decision making (in a dictator game setting).

tures the minimal amount of externalities that is required to rationalise the observed consumption as Pareto efficient. Conveniently, the parameter is situated between zero and one and has a natural degree interpretation. The maximal value of unity means that we can rationalise behaviour in terms of purely selfish consumers (i.e. consumers only care about the own consumption), while the minimal value of zero indicates that rationalisation is possible only for consumers that only care for the others' consumption (and not for the own consumption). Thus, higher parameter values generally suggest that behaviour is more consistent with the standard model of selfish behaviour, while lower values reflect a stronger prevalence of externalities in consumption. By varying the parameter value, we can define a whole continuum of models characterised by different degrees of consumption externalities.

Children and externalities. We use our methodology to investigate the degree of selfishness/externalities of children's joint consumption behaviour.

First of all, we focus on consumption decisions by children. By studying children of different ages, we can examine how selfishness evolves with age. It seems reasonable to argue that the effects of age on pro-social behaviour are more prominent for children than adults. The cognitive developments in children may well be related to significant changes in pro-social behaviour (for an overview of the literature, see e.g. Fehr and Schmidt (1999)). As such, one can expect substantial heterogeneity in selfishness across children of different ages. This is interesting for illustrative purposes because our methodology can perfectly take this heterogeneity into account. Similarly, we can assess the impact of friendship by considering joint consumption decisions of children with various degrees of friendship. We also investigate whether selfishness depends on gender. There is no clear consensus in the literature on how these different variables relate to selfish behaviour. At a more general level, insight in the selfishness of children provides useful information for parents, caretakers and teachers. It determines the extent to which caretakers should guide the distribution of resources among children.

Second, we analyse collective consumption decisions. In this way, we generalise the results by Andreoni and Miller (2002), Fisman et al. (2007) and Cox et al. (2008) who used revealed preference axioms to investigate individual choices in a modified dictator game. These authors invited each individual respondent to allocate money among oneself and a hypothetical counterparty. In our study, we let the children face a real decision-maker, with whom to reach consensus on the consumption of the goods. This increases the external validity of our results. In many settings, children (i.e. siblings, friends, classmates) collectively decide on which activities to engage in, on how to allocate toys or candy, etc. Otherwise stated, we investigate externalities in situations where children are not necessarily dictators. This study also complements the literature on individual rationality in children. Although there has been some research on the individual choices of children (see e.g. Harbaugh et al. (2001) and Bruyneel et al. (2012a)) there is little research on collective decision-making. On the one hand, one could expect low-quality collective decisions when children are not individually competent. On the other hand, given that there is substantial variation in the (individual) rationality of children, it is also possible that the children complement each other, in terms of competent decision-making, when choosing the 'joint' bundle. However, the measurement of decision-making quality depends on the correct specification of children's (other-regarding) preferences, which is all the more relevant in a collective setting.

Finally, two remarks are in order.

Observational data on joint consumption decisions made by children are typically not available. We therefore designed a laboratory experiment that is specially tailored to obtain the data required for our revealed preference methodology. In particular, we first randomly assigned the children that participate to our experiment into dyads. Subsequently, we invited these dyads to jointly choose a series of consumption bundles composed of three commodities (grapes, mandarins and letter biscuits). Once these bundles had been selected, we also registered the associated intra-dyad allocations of the quantities, which gives us all the necessary information to identify our selfishness parameter for the consumption choices that are made.

The interpretation of the selfishness parameter is important, too. Lowering selfishness and hence allowing for externalities is not restricted to implementing 'caring' preferences in the Beckerian sense. It is well known that the so called 'caring' model - with total utility of one individual depending on the egoistic utility of oneself and the egoistic utility of the other - is empirically equivalent to the egoistic version of the collective model (see e.g. Chiappori (1992)). However, this caring model provides a rather narrow definition of altruism because it assumes that the marginal rate of substitution between individually consumed goods is independent of the goods consumed by the other. Especially in a context with children, it seems hard to defend this assumption. After all, it is likely that children directly compare the quantities consumed per commodity. The current framework allows us to investigate whether the egoistic model (and hence the empirically equivalent 'caring' model with the restrictive assumption on externalities) satisfactorily describes the children's decisions. By decreasing our 'selfishness' parameter, we allow for much more general forms of externalities and interdependent marginal rates of substitution.

The remainder of the paper unfolds as follows. Section 3.2 sets out our revealed preference methodology to measure the degree of selfishness/externalities. Section 3.3 presents our experimental design and the results of our empirical application. Section 3.4 concludes.

3.2 Group consumption with non-selfish individuals

To set the stage, we first present the cooperative model under the assumption of selfish group members. Then, we introduce the more general model with non-selfish consumers (i.e. consumption externalities). We show that willingness-to-pay for the other's consumption is captured by personalised prices, and this enables us to define an intuitive selfishness parameter. We conclude this section by discussing some empirical issues related to the practical application of our revealed preference methodology.

Before we can present our models, we first need to specify the type of data that we have in mind when applying our methodology. Our application in Section 3.3 contains information on dyads' consumption behaviour.[5] We have a separate consumption data set for every single dyad, which contains the observed consumption choices for a series of decision situations. Formally, this set takes the form $S = \{(\mathbf{p}_t; \mathbf{q}_t^1, \mathbf{q}_t^2); t = 1, ..., T\}$ and consists of price vectors $\mathbf{p}_t \in \mathbb{R}^n_{++}$ and quantity vectors $\mathbf{q}_t^m \in \mathbb{R}^n_+$ for every observed decision situation t. Each vector $\mathbf{q}_t^m$ represents the quantities of all goods allocated to individual m ($m = 1, 2$).

Throughout this section, we assume that budget sets are linear. In Section 3.3 we briefly discuss the implications of these models for a finite choice set setting.

3.2.1 Selfish individuals

The specific feature of selfish consumer behaviour is that individual utilities are independent of others' consumption. Formally, in our dyad setting each member m has a utility function $U^m(\mathbf{q}^m)$ that only varies with the own consumption $\mathbf{q}^m$. We assume that utility functions are well-behaved (i.e. continuous, monotone and concave). Then, we get the following definition of rational cooperative (i.e. Pareto efficient) consumption behaviour under selfishness.

Definition 3.1. Consider a data set $S = \{(\mathbf{p}_t; \mathbf{q}_t^1, \mathbf{q}_t^2); t = 1, ..., T\}$. A pair of utility functions U^1 and U^2 provides a cooperative rationalisation under selfishness of S if and only if, for each observation $t = 1, ..., T$, there exist Pareto weights $\mu_t^1, \mu_t^2 \in \mathbb{R}_{++}$ such that

[5] We note that it is fairly easy to extend our following methodology towards settings with more than two group members.

$\mu_t^1 U^1\left(\mathbf{q}_t^1\right) + \mu_t^2 U^2\left(\mathbf{q}_t^2\right)$ equals

$$\max_{(\mathbf{z}^1,\,\mathbf{z}^2)\in\left(\mathbb{R}_+^n\right)^2} \mu_t^1 U^1\left(\mathbf{z}^1\right) + \mu_t^2 U^2\left(\mathbf{z}^2\right)$$

$$s.t.$$

$$\mathbf{p}_t'\left(\mathbf{z}^1 + \mathbf{z}^2\right) \leq \mathbf{p}_t'(\mathbf{q}_t^1 + \mathbf{q}_t^2).$$

Thus, Pareto efficiency requires that the dyad's consumption behaviour can be represented as if it maximises a weighted sum of the individual utility functions, subject to the dyad's budget constraint (with the dyad's budget equal to $\mathbf{p}_t'(\mathbf{q}_t^1 + \mathbf{q}_t^2)$). We remark that the individual Pareto weights $\mu_t^1, \mu_t^2 \in \mathbb{R}_{++}$ are allowed to vary across the observations t. The implication is that the 'bargaining power' of a particular individual need not be constant but can depend on the specific decision situation at hand (defined by prices $\mathbf{p}_t$ and budget $\mathbf{p}_t'(\mathbf{q}_t^1 + \mathbf{q}_t^2)$).

Our revealed preference characterisation of rational cooperative behaviour uses the concept GARP (Generalised Axiom of Revealed Preference), which is discussed in Section 2.2. As shown by Varian (1982), consistency with GARP guarantees the existence of an individual utility function U^m that is consistent with the individual m's choices captured by the subset $S^m = \{(\mathbf{p}_t; \mathbf{q}_t^m); t = 1, ..., T\}$. That is, every observed choice $\mathbf{q}_t^m$ maximises this utility function U^m subject to the budget constraint defined by the prices $\mathbf{p}_t$ and the budget $\mathbf{p}_t\mathbf{q}_t^m$.

We can then present the revealed preference characterisation of rational cooperative behaviour with selfish dyad members (see Cherchye et al. (2011a) for a formal proof).

Proposition 3.2. Let $S = \{(\mathbf{p}_t; \mathbf{q}_t^1, \mathbf{q}_t^2); t = 1, ..., T\}$ be a set of observations. The following statements are equivalent:

1. There exists a pair of utility functions U^1 and U^2 that provide a cooperative rationalisation under selfishness of S.

2. The subsets $S^1 = \{(\mathbf{p}_t; \mathbf{q}_t^1); t = 1, ..., T\}$ and $S^2 = \{(\mathbf{p}_t; \mathbf{q}_t^2); t = 1, ..., T\}$ are both consistent with GARP.

Varian (1982) presented a combinatorial test of GARP. More recently, Cherchye et al. (2011a) have shown that the GARP conditions in Proposition 3.2 can also be verified by solving a linear programming problem with binary integer variables. A similar programming problem can also be used to verify the revealed preference conditions in the following Proposition 3.4. For the sake of compactness, and because the analogy with the set-up in Cherchye et al. (2011a) is fairly straightforward, we will not explicitly state this problem in the current paper.

3.2.2 Non-selfish individuals

Non-selfish consumers also care about the other's consumption, which we capture by the utility functions $U^1\left(\mathbf{q}^1, \mathbf{q}^2\right)$ and $U^2\left(\mathbf{q}^1, \mathbf{q}^2\right)$. In our set-up we only consider positive consumption externalities, which means that the functions U^1 and U^2 are increasing in their arguments. Given our particular research question, we use a definition of rational cooperative behaviour that allows for different degrees of externalities. In particular, we capture the degree of selfishness by means of parameters ε and θ. We will explain the meaning of these parameters in more detail below.

Definition 3.3. Consider a data set $S = \{(\mathbf{p}_t; \mathbf{q}_t^1, \mathbf{q}_t^2); t = 1, ..., T\}$. A pair of utility functions U^1 and U^2 provides a cooperative rationalisation under $\theta-$selfishness of S if and only if, for each observation $t = 1, ..., T$, there exist Pareto weights $\mu_t^1, \mu_t^2 \in \mathbb{R}_{++}$ such that

$\mu_t^1 U^1\left(\mathbf{q}_t^1, \mathbf{q}_t^2\right) + \mu_t^2 U^2\left(\mathbf{q}_t^1, \mathbf{q}_t^2\right)$ equals[6]

$$\max_{(\mathbf{z}^1,\,\mathbf{z}^2)\in\left(\mathbb{R}_+^n\right)^2} \mu_t^1 U^1\left(\mathbf{z}^1, \mathbf{z}^2\right) + \mu_t^2 U^2\left(\mathbf{z}^1, \mathbf{z}^2\right)$$

$$s.t.$$

$$\varepsilon = \frac{\theta}{1-\theta},$$

$$\mathbf{p}_t'\left(\mathbf{z}^1 + \mathbf{z}^2\right) \leq \mathbf{p}_t'(\mathbf{q}_t^1 + \mathbf{q}_t^2),$$

$$\varepsilon \leq \frac{\mu_t^1 \frac{\partial U^1}{\partial z_j^1}}{\mu_t^2 \frac{\partial U^2}{\partial z_j^1}} \text{ and } \varepsilon \leq \frac{\mu_t^2 \frac{\partial U^2}{\partial z_j^2}}{\mu_t^1 \frac{\partial U^1}{\partial z_j^2}} \quad \text{with } j = 1, \ldots, n. \tag{3.1}$$

In this definition, the parameter ε relates the marginal willingness-to-pay of member m for his/her own consumption to the one of the other member n ($n \neq m$) for the same consumption. It defines a lower bound on the marginal rate of substitution for every good j between own utility and the utility of the other person. Intuitively, if externalities are prevalent, the marginal willingness-to-pay for the other's consumption will be high, which implies that the data can be rationalised only for a low value of ε. Generally, by varying the value of ε we obtain rationalisation conditions for different degrees of selfishness. We illustrate this by considering the two polar cases. First, when $\varepsilon = 0$, Conditions (3.1) impose no additional restrictions on the optimisation problem. In other words, externalities may be very large. Next, for $\varepsilon \to \infty$ we get exactly the same rationalisation condition as in Definition 3.1, which implies purely selfish dyad members. More generally, greater values of ε correspond to more selfish consumer behaviour. Conveniently, by using the parameter θ we can also derive revealed preference conditions for cooperative rational behaviour that are linear in unknowns, which makes them easy to verify in practice.

To derive a revealed preference test for consistency with Definition 3.3 we first derive the first-order conditions associated with this definition:

[6] z_j^m denotes the j-th component of the vector $\mathbf{z}^m$.

$$\begin{aligned}
\mu_t^1 \frac{\partial U^1}{\partial \mathbf{q}_t^1} + \mu_t^2 \frac{\partial U^2}{\partial \mathbf{q}_t^1} &\leq \lambda_t \mathbf{p}_t, \\
\mu_t^2 \frac{\partial U^2}{\partial \mathbf{q}_t^2} + \mu_t^1 \frac{\partial U^1}{\partial \mathbf{q}_t^2} &\leq \lambda_t \mathbf{p}_t,
\end{aligned}$$

We define the personalised prices

$$\begin{aligned}
\mathbf{p}_t^{1,2} &= \frac{\mu_t^1}{\lambda_t} \frac{\partial U^1}{\partial \mathbf{q}_t^2}, \mathbf{p}_t^{2,1} = \frac{\mu_t^2}{\lambda_t} \frac{\partial U^2}{\partial \mathbf{q}_t^1}, \\
\mathbf{p}_t^{1,1} &= \mathbf{p}_t - \frac{\mu_t^2}{\lambda_t} \frac{\partial U^2}{\partial \mathbf{q}_t^1} \geq \frac{\mu_t^1}{\lambda_t} \frac{\partial U^1}{\partial \mathbf{q}_t^1}, \\
\mathbf{p}_t^{2,2} &= \mathbf{p}_t - \frac{\mu_t^1}{\lambda_t} \frac{\partial U^1}{\partial \mathbf{q}_t^2} \geq \frac{\mu_t^2}{\lambda_t} \frac{\partial U^2}{\partial \mathbf{q}_t^2}.
\end{aligned}$$

where λ_t is the Lagrange multiplier associated with the dyad's optimisation problem in decision situation t (i.e. the marginal value of income). Intuitively, these personalised prices denote the marginal willingness-to-pay for the own and the other's consumption, respectively. Next, each dyad member i decides on own consumption and the other's consumption conditional on his or her budget $\mathbf{p}_t^{i,i\prime}\mathbf{q}_t^i + \mathbf{p}_t^{i,j\prime}\mathbf{q}_t^j$ and his or her shadow prices $\mathbf{p}_t^{i,i}$ and $\mathbf{p}_t^{i,j}$. It is required that these decisions are rational, i.e. consistent with GARP. Finally, the shadow prices $\mathbf{p}_t^{i,i}$ and $\mathbf{p}_t^{j,j}$ are bounded from below by the selfishness parameter θ. This corresponds to Condition (3.1) which restricts the (relative) marginal willingness to pay for own consumption.

Using these concepts, we can state the next result, which generalises Proposition 3.2. (Appendix 3.A contains our proof.)

Proposition 3.4. Let $S = \{(\mathbf{p}_t; \mathbf{q}_t^1, \mathbf{q}_t^2); t = 1, ..., T\}$ be a set of observations. The following statements are equivalent:

1. There exists a pair of utility functions U^1 and U^2 that provide a cooperative rational-

isation under $\theta-$selfishness of S.

2. For all $t = 1, ..., T$, there exist price vectors $\mathbf{p}_t^{1,1}, \mathbf{p}_t^{1,2}, \mathbf{p}_t^{2,1}$ and $\mathbf{p}_t^{2,2} \in \mathbb{R}_+^n$ such that

    (a) the subsets $S^1 = \{(\mathbf{p}_t^{1,1}, \mathbf{p}_t^{1,2}; \mathbf{q}_t^1, \mathbf{q}_t^2); t = 1, ..., T\}$ and
    $S^2 = \{(\mathbf{p}_t^{2,1}, \mathbf{p}_t^{2,2}; \mathbf{q}_t^1, \mathbf{q}_t^2); t = 1, ..., T\}$ both satisfy GARP;

    (b) $\mathbf{p}_t^{1,1} + \mathbf{p}_t^{2,1} = \mathbf{p}_t = \mathbf{p}_t^{1,2} + \mathbf{p}_t^{2,2}$;

    (c) $\mathbf{p}_t^{1,1} \geq \theta \mathbf{p}_t$ and $\mathbf{p}_t^{2,2} \geq \theta \mathbf{p}_t$.

Condition (a) imposes consistency with GARP on the individual subsets S^1 and S^2. Different from Proposition 3.2, these conditions are now expressed in terms of the personalised prices $\mathbf{p}_t^{m,m}$ and $\mathbf{p}_t^{m,n}$ (with $m, n = 1, 2$ and $m \neq n$). Next, Condition (b) states that these personalised prices must add up (over the dyad members) to the observed prices $\mathbf{p}_t$. This condition follows from our assumption that dyads act cooperatively, which means that they achieve Pareto efficient allocations. Actually, the adding up condition also implies that personalised prices can be interpreted as Lindahl prices associated with the Pareto efficient provision of public goods. This corresponds to the fact that private goods with externalities effectively get a public good character.

Condition (c) includes our selfishness parameter θ. It follows that θ measures the fraction of the value of each member m's consumption bundle that (s)he 'finances' him-/herself. As such, this measure has an appealing monetary interpretation. If $\theta = 1$, each member fully pays for her own private consumption, i.e. there are no externalities and behaviour can be rationalised as purely selfish. We then get exactly the conditions for a rationalisation under selfishness that we stated in Proposition 3.2. Lower values of θ enable stronger externalities. In the extreme case with $\theta = 0$, we allow for the possibility that m's consumption is fully financed by the other member n, which means member m does not contribute to his/her own consumption at all.

At this point, it is worth noting that θ puts a lower bound on the monetary contribution of each member for his/her own consumption. This means that when some data set is rationalisable with $\theta-$selfishness, it is also rationalisable for all lower levels of selfishness.

Finally, note that this selfishness parameter θ is not independent of the observed intra-dyad allocation, i.e. the sharing of resources. More importantly, a more equal distribution of resources does not necessarily correspond to a lower degree of selfishness. This is apparent in the following example.

Example 3.5. Consider a situation with $|T| = 2$ observations and $|N| = 2$ goods, scalar x, quantity vectors $\mathbf{q}_1$ and $\mathbf{q}_2$ and the corresponding price vectors $\mathbf{p}_1$ and $\mathbf{p}_2$:

$$\begin{aligned}
\mathbf{q}_1 &= \begin{bmatrix} x & 0 \end{bmatrix}' \text{ and } \mathbf{q}_2 = \begin{bmatrix} 0 & x \end{bmatrix}', \\
\mathbf{p}_1 &= \begin{bmatrix} 2 & 1 \end{bmatrix}' \text{ and } \mathbf{p}_2 = \begin{bmatrix} 1 & 2 \end{bmatrix}'.
\end{aligned}$$

Moreover, consider an allocation rule $\delta \in [0, 1]$ such that the intra-dyad allocation of the goods to dyad member A ($\mathbf{q}_1^A$ and $\mathbf{q}_2^A$) and dyad member B ($\mathbf{q}_1^B$ and $\mathbf{q}_2^B$) equals

$$\begin{aligned}
\mathbf{q}_1^A &= \begin{bmatrix} \delta x & 0 \end{bmatrix}' \text{ and } \mathbf{q}_1^B = \begin{bmatrix} (1-\delta)x & 0 \end{bmatrix}', \\
\mathbf{q}_2^A &= \begin{bmatrix} 0 & \delta x \end{bmatrix}' \text{ and } \mathbf{q}_2^B = \begin{bmatrix} 0 & (1-\delta)x \end{bmatrix}'.
\end{aligned}$$

It is easy to see that when $\delta \neq 0$ then $S^A = \{(\mathbf{p}_t; \mathbf{q}_t^A); t = 1, 2\}$ violates GARP. Likewise, when $\delta \neq 1$ then $S^B = \{(\mathbf{p}_t; \mathbf{q}_t^B); t = 1, 2\}$ violates GARP. There is no value of δ which makes the observations consistent with the egoistic model ($\theta = 1$) of collective rationality. However, we can identify the maximum $\bar{\theta}$ that rationalises the data[7]. In order to rationalise the data when $\delta = 1$, we need $\theta \leq 0.5$. In order to rationalise the data when $\delta = 0.75$, we need $\theta \leq 4/7$. Finally, to rationalise the data when the goods are distributed

[7]The exact formula to compute the maximum level of selfishness for this (simple) example is

equally ($\delta = 0.5$), we find that $\theta \leq 2/3$. Equality of the intra-dyad allocation is related to more selfish behaviour in this example. The intuition behind these results is simple. The dyad in this example spends the total budget on the most expensive good in each observation. This can only be rational if the preferences of dyad members are very heterogeneous and bargaining power shifts substantially. Say, for instance, that member A has a strong preference for good 1, B a strong preference for good 2, and that A has a stronger bargaining position in observation 1. This explains why the consumption of good 1 in observation 1 is important. Likewise, B has a stronger bargaining position in observation 2. However, this is not reflected in the intra-dyad allocation of goods when δ is large, unless member B has strong preferences for the consumption of good 2 by member A. As a result, the more unequal the intra-dyad allocation (i.e. the larger $|\delta - 0.5|$) the lower the maximum possible level of selfishness.

3.2.3 Empirical concerns

Practical applications of revealed preference conditions like the ones in Propositions 3.2 and 3.4 typically raise two empirical concerns: the possibility of optimisation error (i.e. behaviour may be close to but not exactly optimising) and the issue of discriminatory power (i.e. the empirical stringency of the optimisation conditions). We next discuss each of these issues in more detail, and indicate how we will deal with them in our empirical application in Section 3.3.

Optimisation error. Our revealed preference conditions are based on the assumption of 'exactly' optimising behaviour. Obviously, this may often seem like an overly restrictive

---

$$\begin{aligned}\bar{\theta} &= \min\left(\frac{p_2^1}{\delta p_2^2 + (1-\delta)p_2^1}, \frac{p_2^1}{(1-\delta)p_2^2 + \delta p_2^1}\right) \\ &= \min\left(\frac{1}{\delta 2 + (1-\delta)}, \frac{1}{(1-\delta)2 + \delta}\right).\end{aligned}$$

assumption, which holds all the more true for our following application to consumption decisions of children. Irrespective of their degree of selfishness, individuals are not always able, or willing, to behave as fully rational (i.e. utility maximising) homines economici. However, it may well be that their decisions are close to being rational (while not exactly rational). Putting it differently, it is often realistic to allow for a certain (small) degree of optimisation error. Our application in Section 3.3 will use an extension of our methodology that accounts for such optimisation error.

Following Afriat (1972) and Varian (1990), we account for 'nearly' optimising behaviour by applying Afriat's Critical Cost Efficiency Index. My General Introduction contains a discussion of this index and the corresponding GARP characterisation (Definition 0.2).

As explained in detail by Choi, Kariv, Müller, and Silverman (2014), this index (denoted e) actually captures how much the individual's budget needs to be reduced in order to make the observed choices consistent with GARP. Or conversely, the fraction $1 - e$ measures the fraction of the budget that has been wasted due to irrational consumption decisions.

Pass rate, power and predictive success. We have presented a continuum of models where lower values of θ allow for more consumption externalities. Thus, by construction we will have that lower θ-values lead to less restrictive consumption models, which makes it easier to pass the corresponding revealed preference conditions. To account for this trade-off between economic realism (i.e. permit deviations from purely selfish behaviour) and restrictiveness, a fair comparison of models with different θ-values should simultaneously account for both their empirical fit (i.e. whether or not the data satisfy the associated rationalisation conditions) and their discriminatory power (i.e. the extent to which these rationalisation conditions can effectively identify irrational behaviour). Ideally, a behavioural model combines a good empirical fit with high discriminatory power. To capture this idea, our empirical analysis in Section 3.3 will assess alternative models in terms of their 'predictive success', which combines empirical fit and discriminatory power in a single metric.

For a particular behavioural model (defined by a specific optimisation parameter e and selfishness parameter θ), we compute the fraction of observed data sets that satisfy the corresponding rationalisation conditions. We call this fraction our pass rate. Its interpretation is discussed in detail in Chapter 2.

To measure the discriminatory power of a behavioural model, we make use of Bronars (1987)'s power index. The computation is similar to the approach discussed in Chapter 2, but there is an additional step. In the first stage, we draw bundles from the uniform distribution of bundles in the respective choice sets. In the second stage, we divide the goods among the dyad members by using (random) shares that vary across the goods. Finally, we obtain a new data set containing T simulated choices (and the corresponding intra-dyad allocation). We repeat this procedure 5000 times, which thus defines 5000 sets of T 'irrational' consumption choices. Bronars' power index equals one minus the fraction of these simulated data sets that pass the rationalisation conditions under evaluation.

Our measure of predictive success simultaneously includes pass rate and power. As such, it takes into account both the potential of a model to describe the observed behaviour (captured by the pass rate) as well as the potential to detect irrational behaviour (captured by Bronars' power index).

$$\text{Predictive success} = \text{Pass rate} - [1 - \text{Power}].$$

Similar to before, a value close to one indicates a model with approximately perfect discriminatory power and fit, i.e. the best possible scenario. This means that (almost) all data pass the rationality test, even though the test effectively detects (almost) any deviating (i.e. irrational) behaviour. By contrast, a value close to minus one implies a model with almost no discriminatory power and a very bad fit, i.e. the worst possible scenario. Finally, a value of zero corresponds to a model with a pass rate for the observed behaviour that exactly equals the expected pass rate if behaviour were irrational. Essentially, this means that the rationality test does not allow for distinguishing observed behaviour from irrational behaviour.

3.3 Joint decisions of children

Before we present our empirical results, we first explain our experimental design. In doing so, we will also motivate the empirical questions that we consider further on, with references to the relevant literature. Subsequently, we discuss the main results of our empirical analysis. In particular, we find strong evidence that children's joint consumption behaviour is systematically characterised by consumption externalities (i.e. non-selfish behaviour). Interestingly, we also observe that these externalities bear particular relations to age, gender and the degree of friendship.

3.3.1 Experimental design

Respondents. We collected our data at four different schools. The selection of classes and schools in the sample is presented in Table 3.1. Our sample contains a total of 100 children, who belong to 3 different age categories: 42 from kindergarten, 24 from third grade and 34 from sixth grade.

	kindergarten	third grade	sixth grade
School I	1 class (8)	1 class (8)	1 class (6)
School II	1 class (7)	1 class (4)	1 class (11)
School III	1 class (6)	0 classes	0 classes

Table 3.1: Information on schools and classes (number of dyads per class)

Table 3.2 gives summary information for our sample in terms of gender composition and the degree of friendship (explained below). In what follows, we discuss the construction of our sample in more detail, and use this to position our following empirical analysis in the existing literature.

First of all, our sample allows us to link selfishness to children's age. There is some evidence that people in early childhood (children aged less than 5) are less altruistic (see, for example, Eisenberg et al. (2007) for a literature review on the development of prosocial behaviour) and more likely to be driven by self-interest (see, for example, Damon (1980) on positive-justice). However, this does not automatically imply a stable and decreasing relationship between age and selfishness. On the one hand, Côté et al. (2002) found support for inter-individual stability in prosocial behaviour. Similarly, Gummerum et al. (2008) did not find significant age effects on individual allocations in a dictator game. On the other hand, there is also evidence that young school children sometimes act less selfishly. See, for example, Murnighan and Saxon (1998) and Harbaugh et al. (2003), who found that younger children are more likely to accept smaller offers in ultimatum games, or Damon (1980), who found that children from 5 to 7 years of age frequently selected equal rewards in order to avoid conflict. By the age of 5, children tend to have egalitarian preferences and select outcomes that distribute pay-offs equally.

In this respect, a particularly interesting study is the one of Fehr et al. (2013). These authors argue that, even though altruism generally increases with age, preferences for egalitarianism seem to peak around the age of eight years (which corresponds to our group of third graders). Beyond this age, the increasing influence of efficiency-considerations and strategic behaviour may countervail fairness-considerations. In a similar vein, it is claimed that the positive effects of a more prosocial orientation are offset by increasing levels of competitiveness as children grow older. Kagan and Madsen (1972) and Toda et al. (1978), for instance, have shown that the level of competition between children increases as a function of age.

Summarising, we may safely conclude that the literature does not show a clear consensus on the relationship between age and selfish preferences. This directly provides a particular motivation for our own empirical application. We deliver empirical input to the debate by considering selfishness in the specific context of children's group consumption decisions.

For each separate age category, we randomly organised the children into dyads (i.e. two-member groups), which we then invited to make 9 consumption choices. This resulted in 50 dyads and obtained information on 450 (= 50 · 9) joint decisions. We registered the gender composition of each dyad. There are 19 female dyads, 12 male dyads and 19 dyads consisting of one boy and one girl. Eisenberg et al. (2007) argue that girls are more prosocial than boys. Moreover, girls tend to be somewhat less competitive. Similar to before, our analysis will allow us to investigate this further in a specific consumption context.

Finally, we also registered the intensity of the dyad members' relationship outside the experiment. In particular, we asked the children to label their relationship with respect to the other dyad member as '(very) strong friendship' or 'weak (or no) friendship'. According to Eisenberg et al. (2007), the literature suggests that children are more likely to share with friends than with less liked peers (see also Buhrmester et al. (1992) and Pilgrim and Rueda-Riedle (2002)). We will investigate this effect in a group consumption context. In a sense, a minimal requirement for our measure of selfishness to be a sensible one is that it bears a negative relationship to self-reported friendship.

		gender	
		boy	girl
grade	kindergarten	4 / 12 / 2	11 / 13 / 0
	thirdgrade	4 / 3 / 2	8 / 7 / 0
	sixthgrade	7 / 8 / 1	5 / 12 / 1

Table 3.2: Summary statistics on sample composition (x/y/z: x children who indicate (very) strong friendship with their dyad partner, y children who indicate weak (or no) friendship with their dyad partner, z children with missing values on the relationship)

Design. The experimental design is almost identical to the set-up described in Section 1.2. In the same way, we presented discrete choice sets (containing 7 different combinations of grapes, mandarins and letter biscuits) to the respondents. The implicit prices are also identical to the experiment in Chapter 1 but the implicit budgets are 24 (instead of 12).

The experiment under consideration proceeded in two basic steps. In a first step, each dyad of children was asked to select one out of seven possible commodity bundles for the given (implicit) price and budget regimes. The children could take as much time as they wanted to take their joint decisions. In a second step, and in view of our following assessment of externalities, we asked each dyad to define individual shares of the joint consumption bundle that had been chosen, which makes that we perfectly observe the shares of the (implicit) dyad budget allocated to each individual member.

	Obs	Mean	Std. Dev.	Min	Max	% Equal
all	50	0.064	0.064	0.001	0.29	56
kindergarten	21	0.099	0.081	0.004	0.29	33.33
third grade	12	0.029	0.018	0.001	0.06	91.67
sixth grade	17	0.046	0.032	0.004	0.122	58.82
weak friendship	30	0.077	0.076	0.001	0.29	46.67
strong friendship	17	0.048	0.034	0.001	0.122	64.71
two girls	19	0.067	0.065	0.001	0.29	52.63
mixed	19	0.063	0.055	0.001	0.18	52.63
two boys	12	0.062	0.078	0.004	0.29	66.67

Table 3.3: Intra-dyad budget sharing: summary statistics for the absolute intra-dyad difference in allocated budgets. The final column gives the percentage of dyads in which the implicit budget is shared equally (with absolute differences less than 5 per cent).

Table 3.3 reports summary statistics on the absolute intra-dyad differences between individual budget shares, which provides some basic insight into the intra-dyad sharing of resources. The table also gives the proportion of dyads that apply (close to) equal resource sharing (i.e. intra-dyad difference between individual resource shares amounts to less than 5 percent of the available budget). We find that, on average, the resources are shared fairly equally. The mean absolute intra-dyad difference in shares amounts to 6.4 percent. Interestingly, the difference is smallest for dyads containing third graders, while it is largest for dyads with kindergarten respondents. Similarly, we observe that sharing is more equal when children have a strong friendship relationship with their partner. Finally, the gender composition does not seem to have a strong impact on the resource sharing pattern. Importantly, the goal of our empirical analysis goes beyond simply describing the sharing of resources. We want to investigate if externalities in consumption impact the decision processes that underlie the patterns summarised in Table 3.3. And, if so, we want to quantify the willingness-to-pay for these externalities, and relate the associated degree of selfishness to the children characteristics reported in Table 3.2.

Finally, notice that the choice sets are finite. From Chapter 2, on the one hand, it is clear that the standard revealed preference test for individual rationality (GARP) is overly strong when choice sets are finite. In a collective setting, on the other hand, the finiteness is partly neutralised by the fact that dyad members can allocate the chosen quantities continuously among each other. As such, individual i's (personalised) budget set is no longer purely finite even when the dyad-level choice set is finite. Nonetheless, the finiteness of the choice sets may cause specific bundles to be unavailable even if they are located in the lower half-space of the hyperplane defined by the budget $\mathbf{p}_t^{i,i\prime}\mathbf{q}_t^i + \mathbf{p}_t^{i,j\prime}\mathbf{q}_t^j$ and shadow prices $\mathbf{p}_t^{i,i}$ and $\mathbf{p}_t^{i,j}$. The proposed test in Section 3.2 should therefore be interpreted as a sufficient (but not necessary) test for rationality[8]. This also implies that we identify an *upper bound* on the minimum required marginal willingness to pay for others' consumption (or equivalently, a *lower bound* on the maximum possible level of selfishness) in Subsection 3.3.3.

3.3.2 Consumption with or without externalities?

We begin our analysis by evaluating the empirical performance of the cooperative consumption model that we characterised in Proposition 3.4, for alternative values of the selfishness parameter θ. We recall that this parameter ranges from zero to one, with $\theta = 1$ indicating purely selfish behaviour and $\theta = 0$ defining a least restrictive model that also accounts for the (opposite) scenario in which dyad members only care about the other's consumption. As explained in Section 3.2, our empirical analysis considers the possibility of optimisation error, which we capture by the parameter e. In what follows, we let this parameter

[8] As an alternative approach, one could take the finiteness in the collective setting into account by building on Proposition 2 in Cherchye et al. (2007), and by replacing the (collective) budget sets in this proposition with finite choice sets. This would lead to a necessary (but not sufficient) test for consistency. Moreover, the resulting characterisation would no longer allow us to formulate selfishness in monetary terms. Shadow prices are not identified. We therefore believe that the original (sufficient) test is more valuable to address our research question.

vary between 1 (i.e. no optimisation error) and 0.9 (i.e. an optimisation error of at most 10 percent).[9] The corresponding pass rates, power results and predictive success values are summarised in Tables 3.4 to 3.6.

Let us first focus on the pass rates in Table 3.4, which shed light on the goodness-of-fit of the different models under study. We find that the choices of all dyads can be rationalised for $\theta = 0$. This should actually not be too surprising because, as explained before, this defines a very permissive model of cooperative consumption behaviour. Next, we also observe that, for any θ, the pass rate increases (slightly) if we allow for some optimisation error (i.e. $e < 1$). Again, this is as expected as lower values for e generally imply less stringent rationalisation conditions.

We next consider the permissiveness of the models, by evaluating their discriminatory power. These results are given in Table 3.5. We find that the standard model with purely selfish consumers and no optimisation error is indeed a very stringent one, as it is characterised by a discriminatory power of about 0.971. In other words, (simulated) irrational behaviour passes the associated rationalisation conditions in less than 3 percent of the cases. Generally, we observe that the rationalisation conditions become more permissive if we leave more room for consumption externalities (i.e. non-selfish behaviour, captured by lower θ-values). Interestingly, the effect of lowering e (i.e. more optimisation error) is less pronounced. For $\theta = 0$ our power index is only about 5 percent, which signals a very low probability of detecting irrational behaviour. Importantly, however, discriminatory power does remain rather high as long as the degree of non-selfishness is somewhat restricted. For example, for $\theta \geq 0.75$ we find that random behaviour is still diagnosed as irrational in more than 70 percent of the cases, even if we allow for 10 percent optimisation error.

Our predictive success results are reported in Table 3.6. This table presents summary scores for the overall performance of the different model specifications that we assess. For most values of our selfishness parameter (the only exceptions are $\theta = 0.95$ and $\theta = 0$), we

[9] This follows an original suggestion of Varian (1990), who proposed to choose $e = 0.9$ as a cut-off value.

find highest predictive success scores if we allow for some optimisation error (i.e. $e < 1$). This suggests that, in general, children's consumption behaviour is nearly optimising rather than exactly optimising. More interestingly, the highest predictive success score is realised for $e = 0.90$ and $\theta = 0.90$, i.e. 10 percent optimisation error and some room for non-selfish behaviour. For our sample, this model specification has a predictive success of 0.490, which is substantially above the predictive success of any other specification under study. This predictive success score follows from a pass rate of 0.64 (see Table 3.4) and a power index of 0.850 (see Table 3.5).

At a more general level, we conclude from the results in Table 3.6 that children's consumption behaviour is systematically characterised by consumption externalities: consumption models that account for (a limited amount of) non-selfishness have a higher predictive success than the purely selfish model. However, as we discuss next, it will also appear that children are quite heterogeneous in their degree of non-selfishness.

pr	$\theta = 1$	0.95	0.9	0.85	0.8	0.75	0.5	0.25	0
$e =$									
1	0.46	0.5	0.52	0.56	0.58	0.58	0.74	0.94	1
	[0.07]	[0.07]	[0.07]	[0.07]	[0.07]	[0.07]	[0.06]	[0.03]	[0]
0.99	0.46	0.5	0.52	0.56	0.58	0.58	0.76	0.94	1
	[0.07]	[0.07]	[0.07]	[0.07]	[0.07]	[0.07]	[0.06]	[0.03]	[0]
0.95	0.5	0.52	0.56	0.58	0.58	0.64	0.78	0.96	1
	[0.07]	[0.07]	[0.07]	[0.07]	[0.07]	[0.07]	[0.06]	[0.03]	[0]
0.9	0.5	0.54	0.64	0.66	0.66	0.72	0.86	0.96	1
	[0.07]	[0.07]	[0.07]	[0.07]	[0.07]	[0.06]	[0.05]	[0.03]	[0]

Table 3.4: Collective rationality: pass rates, [std.dev]

power	$\theta = 1$	0.95	0.9	0.85	0.8	0.75	0.5	0.25	0
$e =$									
1	0.971	0.949	0.917	0.875	0.825	0.783	0.599	0.285	0.055
0.99	0.969	0.944	0.909	0.865	0.815	0.774	0.591	0.276	0.048
0.95	0.955	0.925	0.882	0.837	0.786	0.745	0.569	0.246	0.048
0.9	0.936	0.895	0.850	0.795	0.753	0.711	0.542	0.208	0.048

Table 3.5: Collective rationality: power

pr succes	$\theta = 1$	0.95	0.9	0.85	0.8	0.75	0.5	0.25	0
$e =$									
1	0.431	0.449	0.437	0.435	0.405	0.363	0.339	0.225	0.055
0.99	0.429	0.444	0.429	0.425	0.395	0.354	0.351	0.216	0.048
0.95	0.455	0.445	0.442	0.417	0.366	0.385	0.349	0.206	0.048
0.9	0.436	0.435	0.490	0.455	0.413	0.431	0.402	0.168	0.048

Table 3.6: Collective rationality: predictive success

3.3.3 Selfishness and child characteristics

Individual selfishness. In order to investigate heterogeneity in selfishness (captured by θ) across children in our sample, we no longer consider a common θ for both members of a given dyad, as in our original formulation of Proposition 3.4. Instead, we define a different θ^m for each dyad member m. The corresponding adaptation of the rationalisation conditions in Proposition 3.4 is immediate.[10]

For a given data set on dyad consumption choices, we maximise the average $\theta^* = (\theta^1 + \theta^2)/2$ subject to the given rationalisation conditions. Basically, this computes (an upper bound on) the minimal degree of non-selfishness (i.e. consumption externalities) that we need to account for in order to rationalise the observed dyad behaviour in terms of the cooperative model (under equal weighting of the dyad members). Lower values of θ^* (and, correspondingly, θ^1 and θ^2) indicate that consistency with cooperative group behaviour requires greater deviations from purely selfish behaviour.

By maximising $\theta^* = (\theta^1 + \theta^2)/2$, we can compute a selfishness parameter θ^m for each different individual m in our sample. Figure 3.1 presents the distribution of this selfishness parameter for the 100 individuals in our experiment. We find that for about 64 percent of the individuals m the value of θ^m equals unity. For the remaining children, we need to account for consumption externalities (i.e. non-selfish preferences) to rationalise the observed consumption behaviour. Actually, we observe a positive density even for θ^m as low as 0.2, which reveals a high degree of non-selfishness. Generally, the distribution pattern in Figure

[10] In terms of the cooperative rationalisation concept in Definition 3.3, this boils down to using ε^m for each m (instead of a common ε).

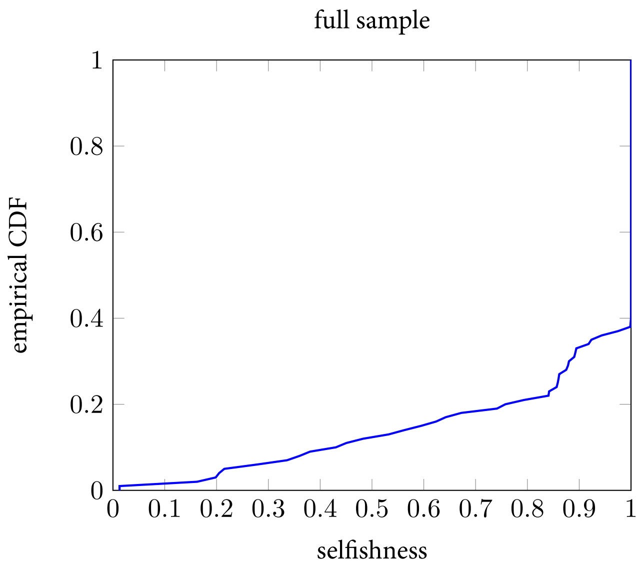


Figure 3.1: Cumulative distribution of the individual selfishness parameter; whole sample

3.1 reveals considerable inter-individual heterogeneity in the degree of selfishness.

In Appendix 3.B, we investigate the sensitivity of the distribution of the selfishness parameter to different degrees of randomness in the data. We find that the distribution is robust to limited levels of noise.

Relation with observed characteristics. We first consider how friendship relates to individual selfishness. In particular, we distinguish between two types of children: children who report a (very) strong friendship with their dyad partner, and children who report a weak (or no) friendship. Our results are displayed in Figure 3.2. The two curves in this figure exhibit a clear first order stochastic dominance relationship, which indicates that the degree of selfishness in behaviour is systematically lower when children make joint consumption decisions with other children who they consider to be strong friends. In this case, children are willing to contribute more to the other's material consumption (i.e. consumption externalities). This is exactly what can be expected from friends, and falls in line with the literature

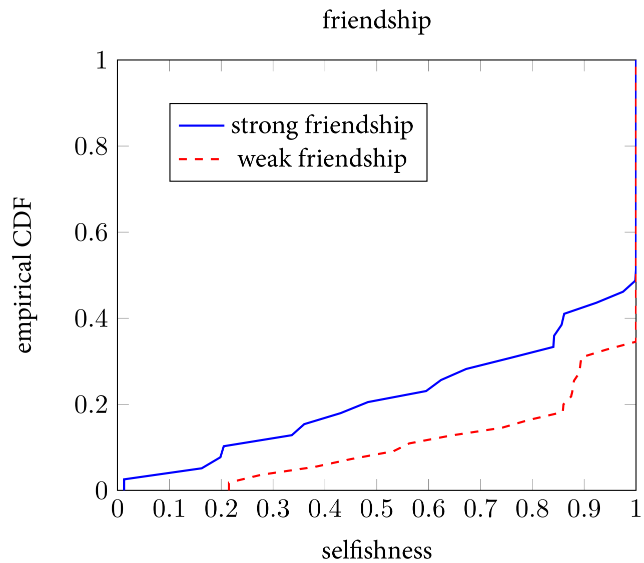


Figure 3.2: Cumulative distribution of the individual selfishness parameter; strong friendship versus weak friendship

(see, for example, Buhrmester et al. (1992), Pilgrim and Rueda-Riedle (2002) and Eisenberg et al. (2007)). In a sense, this also indicates that our methodology effectively does produce a sensible measure of selfishness.

Next, we turn to the gender effect, for which the relevant results are given in Figure 3.3. Interestingly, we get a similar dominance relationship as in Figure 3.2. In this particular case, this provides a clear indication that girls care more for the consumption of others than boys. As discussed above, this falls in line with reported evidence that girls generally do tend to act more prosocially (and less competitive).

Finally, we consider the age effect, for which there appeared to be no clear consensus in the literature. The results are summarised in Figure 3.4. A first observation here is that the θ^m-scores for kindergarten respondents and third graders are roughly similar. Next, we also find that sixth graders are generally more selfish than younger children (both kindergarten children and third graders), who seem to be characterised by larger consumption

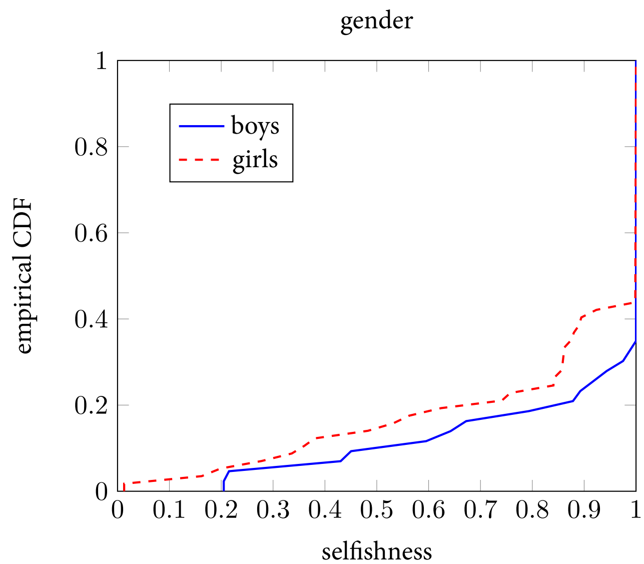


Figure 3.3: Cumulative distribution of the individual selfishness parameter; boys versus girls

externalities.

At first glance, these results may seem to contradict the conclusion of Eisenberg et al. (2007), which indicates a positive relationship between age and prosocial behaviour. In this respect, however, we also recall the study by Fehr et al. (2013), who found that preferences for egalitarianism peak around the age of eight years (i.e. third grade) and decrease beyond this age (which was also partly reflected by the average differences in shares). Moreover, we also argued that incidences of competitiveness between children and strategic behaviour appear to increase with age (see, for example, Kagan and Madsen (1972) and Toda et al. (1978)). As such, our results clearly provide further input to this interesting debate, by focusing on the specific setting of joint consumption decisions. At a more general level, this also nicely motivates the practical usefulness of our methodology.

Statistical significance. To verify the statistical meaning of our above conclusions, we carried out Wilcoxon ranksum tests (or Mann-Whitney U tests). (As an additional robustness

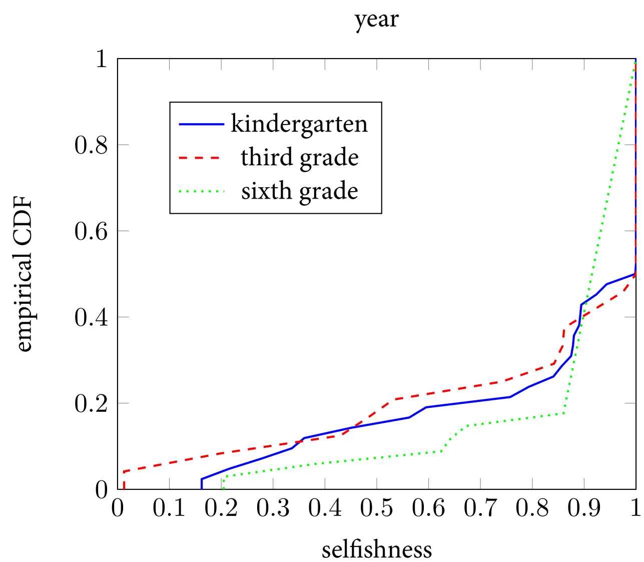


Figure 3.4: Cumulative distribution of the individual selfishness parameter; kindergarten, third grade and sixth grade

check, we also conducted Kolmogorov-Smirnov tests, of which the results are reported in Appendix 3.C.) In these exercises, the null hypothesis is that two populations have the same distribution for our individual selfishness parameter θ^m. Correspondingly, the alternative hypothesis is that one of the populations systematically has higher values for the parameter than the other. The results of our Wilcoxon tests are given in Table 3.7. Interestingly, these test results do confirm that, in general, our above conclusions are statistically robust.

Specifically, for the age effect we find that kindergarten respondents and third graders have a lower rank sum than expected under the null hypothesis, whereas sixth graders have a higher rank sum than expected. Correspondingly, we reject the hypothesis that θ^m is equally distributed for the two groups (i.e. kindergarten respondents and third graders versus sixth graders). Next, the values of θ^m are significantly higher (at the 0.1 level) for children who make consumption decisions with strong friends. Finally, we also observe that boys are more selfish than girls, although the effect here is not strongly significant.

Ranksum		rank sum	expected	P-value
age	0 kind.,thirdg.	3014	3333	
	1 sixthg.	2036	1717	
	Combined	5050	5050	0.0078
friendship	0 weak	2826	2612.5	
	1 strong	1639	1858.5	
	Combined	4465	4465	0.0631
gender	0 girls	2696	2878.5	
	1 boys	2354	2171.5	
	Combined	5050	5050	0.1454

Table 3.7: Ranksum tests

3.4 Conclusion

This paper has both a methodological and an empirical contribution. At the methodological level, we have introduced a revealed preference approach to quantify the willingness-to-pay for the consumption of others. Within the framework of the cooperative (i.e. Pareto efficient) consumption model, we measure willingness-to-pay for others' consumption by evaluating positive consumption externalities in monetary terms. Interestingly, the method allows us to define a selfishness parameter that characterises a continuum of models with varying degrees of consumption externalities.

Next, at the empirical level, we have shown the practical usefulness of our method by an application to consumption choices made by dyads of children. We find that children's consumption decisions are systematically characterised by externalities (i.e. non-selfish). But we also observe that there is substantial heterogeneity across children, which we related to differences in age, gender and degree of friendship between dyad members. For our sample, we found that sixth graders behave more selfishly than third graders and kindergarten children, that boys behave more selfishly than girls (albeit that this effect is not strongly statistically significant), and that children act less selfishly in joint consumption decisions when they have a strong friendship with the other group members.

We see several avenues for further research. At the methodological level, we can extend our revealed preference characterisations to other types of social (or other-regarding) preferences (see, for example, Sobel (2005) for a recent review). For example, we could incorporate negative consumption externalities (including envy) in the analysis. Such an extension would lead to an operational measure that evaluates these negative externalities in monetary terms, which enables similar empirical applications as the one we presented in Section 3.3. In a similar vein, we can also use our revealed preference approach to devise testable implications of more specific types of social preferences (defining particular origins of positive and/or negative externalities). This can be used to investigate whether alternative models are empirically distinguishable from each other in revealed preference terms. And, if so, we can relate the applicability of specific models to the (observable) characteristics of the individuals at hand.

At the empirical level, our application has used data that we collected through a specially designed consumption experiment. This experiment clearly showed the potential of our approach to empirically explore relations between non-selfish behaviour and individual characteristics. In this first study we used only a fairly limited amount of information on individual characteristics (i.e. age, gender and friendship). Obviously, richer data sets can obtain a more detailed analysis of the drivers of positive externalities. For example, this may imply a deeper investigation of the relationship between age and non-selfishness.

Finally, in this study we used experimental data because our focus was on children's consumption. However, our revealed preference methodology can also be used in combination with observational data. For example, an interesting application may identify the degree of selfishness in household consumption, and relate inter-household heterogeneity in our selfishness parameter to specific household (member) characteristics.[11]

[11] Such an application would require observations on the intra-household sharing of consumption. Interestingly, data sets with detailed information on the intra-household consumption allocation are increasingly available. See Cherchye et al. (2012) for a recent example. In this respect, we also refer to Cherchye et al. (2009, 2011a) for empirical studies of household consumption behaviour that make use of revealed preference methods similar to ours.

3.A Proof of Proposition 3.4

Necessity. We show that statement *1* implies statement *2*, i.e. the existence of a pair of utility functions U^1 and U^2 that provide a cooperative rationalisation under $\theta-$selfishness implies that there exist non-negative price vectors $\mathbf{p}_t^{1,1}$, $\mathbf{p}_t^{2,2}$, $\mathbf{p}_t^{1,2}$ and $\mathbf{p}_t^{2,1}$ such that the subsets $S^1 = \{(\mathbf{p}_t^{1,1}, \mathbf{p}_t^{1,2}; \mathbf{q}_t^1, \mathbf{q}_t^2); t = 1, ..., T\}$ and $S^2 = \{(\mathbf{p}_t^{2,1}, \mathbf{p}_t^{2,2}; \mathbf{q}_t^1, \mathbf{q}_t^2); t = 1, ..., T\}$ are both consistent with the GARP and such that the conditions on these price vectors hold.

In a first step, we derive the first-order conditions associated with the optimisation problem in Definition 3.3:

$$\begin{aligned} \mu_t^1 \frac{\partial U^1}{\partial \mathbf{q}_t^1} + \mu_t^2 \frac{\partial U^2}{\partial \mathbf{q}_t^1} &\leq \lambda_t \mathbf{p}_t, \\ \mu_t^2 \frac{\partial U^2}{\partial \mathbf{q}_t^2} + \mu_t^1 \frac{\partial U^1}{\partial \mathbf{q}_t^2} &\leq \lambda_t \mathbf{p}_t, \end{aligned}$$

with $\dfrac{\partial U^m}{\partial \mathbf{q}_t^m}$ and $\dfrac{\partial U^n}{\partial \mathbf{q}_t^m}$ $(m, n = 1, 2, m \neq n)$ the supergradients of the functions U^1 and U^2 with respect to $\mathbf{q}_t^m$, both evaluated at $(\mathbf{q}_t^1, \mathbf{q}_t^2)$. At this point, we can define personalised prices as follows:

$$\begin{aligned} \mathbf{p}_t^{1,2} = \frac{\mu_t^1}{\lambda_t} \frac{\partial U^1}{\partial \mathbf{q}_t^2}, \mathbf{p}_t^{2,1} = \frac{\mu_t^2}{\lambda_t} \frac{\partial U^2}{\partial \mathbf{q}_t^1}, \\ \mathbf{p}_t^{1,1} = \mathbf{p}_t - \frac{\mu_t^2}{\lambda_t} \frac{\partial U^2}{\partial \mathbf{q}_t^1}, \mathbf{p}_t^{2,2} = \mathbf{p}_t - \frac{\mu_t^1}{\lambda_t} \frac{\partial U^1}{\partial \mathbf{q}_t^2}. \end{aligned}$$

This obtains that $\mathbf{p}_t^{1,1} + \mathbf{p}_t^{2,1} = \mathbf{p}_t = \mathbf{p}_t^{1,2} + \mathbf{p}_t^{2,2}$, *which gives Condition 2b*. Moreover, the above shows that

$$\mathbf{p}_t^{1,1} \geq \frac{\mu_t^1}{\lambda_t}\frac{\partial U^1}{\partial \mathbf{q}_t^1}, \mathbf{p}_t^{2,2} \geq \frac{\mu_t^2}{\lambda_t}\frac{\partial U^2}{\partial \mathbf{q}_t^2}.$$

In a second step, we use that the individual utility functions are concave. As such

$$\begin{aligned} U^1(\mathbf{q}_s^1, \mathbf{q}_s^2) - U^1(\mathbf{q}_t^1, \mathbf{q}_t^2) &\leq \frac{\partial U^{1\prime}}{\partial \mathbf{q}_t^1}(\mathbf{q}_s^1 - \mathbf{q}_t^1) + \frac{\partial U^{1\prime}}{\partial \mathbf{q}_t^2}(\mathbf{q}_s^2 - \mathbf{q}_t^2), \\ U^2(\mathbf{q}_s^2, \mathbf{q}_s^1) - U^2(\mathbf{q}_t^2, \mathbf{q}_t^1) &\leq \frac{\partial U^{2\prime}}{\partial \mathbf{q}_t^2}(\mathbf{q}_s^2 - \mathbf{q}_t^2) + \frac{\partial U^{2\prime}}{\partial \mathbf{q}_t^1}(\mathbf{q}_s^1 - \mathbf{q}_t^1). \end{aligned}$$

By taking $\eta_t^m = \dfrac{\lambda_t}{\mu_t^m}$, and given the definitions of $\mathbf{p}_t^{1,1}$, $\mathbf{p}_t^{2,2}$, $\mathbf{p}_t^{1,2}$ and $\mathbf{p}_t^{2,1}$, we then effectively obtain

$$\begin{aligned} U^1(\mathbf{q}_s^1, \mathbf{q}_s^2) - U^1(\mathbf{q}_t^1, \mathbf{q}_t^2) &\leq \eta_t^1 \mathbf{p}_t^{1,1\prime}(\mathbf{q}_s^1 - \mathbf{q}_t^1) + \eta_t^1 \mathbf{p}_t^{1,2\prime}(\mathbf{q}_s^2 - \mathbf{q}_t^2), \\ U^2(\mathbf{q}_s^2, \mathbf{q}_s^1) - U^2(\mathbf{q}_t^2, \mathbf{q}_t^1) &\leq \eta_t^2 \mathbf{p}_t^{2,2\prime}(\mathbf{q}_s^2 - \mathbf{q}_t^2) + \eta_t^2 \mathbf{p}_t^{2,1\prime}(\mathbf{q}_s^1 - \mathbf{q}_t^1). \end{aligned}$$

Taking $U^m(\mathbf{q}_s^m, \mathbf{q}_s^n) = U_s^m$ results exactly into the Afriat inequalities applied to our framework. Varian (1982) proved the equivalence between consistency with the Afriat inequalities and consistency with the GARP. Hence, we have shown that the data set must be such that $S^1 = \{(\mathbf{p}_t^{1,1}, \mathbf{p}_t^{1,2}; \mathbf{q}_t^1, \mathbf{q}_t^2); t = 1, ..., T\}$ and $S^2 = \{(\mathbf{p}_t^{2,1}, \mathbf{p}_t^{2,2}; \mathbf{q}_t^1, \mathbf{q}_t^2); t = 1, ..., T\}$ are both consistent with the GARP. *This gives Condition 2a*.

In a final step, we must take into account that the utility functions U^1 and U^2 were restricted to satisfy

$$\varepsilon \leq \frac{\mu_t^1 \frac{\partial U^1}{\partial z_j^1}}{\mu_t^2 \frac{\partial U^2}{\partial z_j^1}} \text{ and } \varepsilon \leq \frac{\mu_t^2 \frac{\partial U^2}{\partial z_j^2}}{\mu_t^1 \frac{\partial U^1}{\partial z_j^2}} \quad \text{with } j = 1, \ldots, n.$$

Using the above notation, we can rewrite this in terms of personalised prices:

$$\varepsilon \leq \frac{p_{t,j}^{1,1}}{p_{t,j}^{2,1}} \text{ and } \varepsilon \leq \frac{p_{t,j}^{2,2}}{p_{t,j}^{1,2}}.$$

This gives $\mathbf{p}_t^{1,1} \geq \varepsilon \mathbf{p}_t^{2,1}$ and $\mathbf{p}_t^{2,2} \geq \varepsilon \mathbf{p}_t^{1,2}$. Hence $\mathbf{p}_t^{1,1} \geq \varepsilon(\mathbf{p}_t - \mathbf{p}_t^{1,1})$ and $\mathbf{p}_t^{2,2} \geq \varepsilon(\mathbf{p}_t - \mathbf{p}_t^{2,2})$ or $\mathbf{p}_t^{1,1} \geq \dfrac{\varepsilon}{1+\varepsilon}\mathbf{p}_t$ and $\mathbf{p}_t^{2,2} \geq \dfrac{\varepsilon}{1+\varepsilon}\mathbf{p}_t$. Given the definition of the selfishness parameter, $\theta = \dfrac{\varepsilon}{1+\varepsilon}$, we thus obtain $\mathbf{p}_t^{1,1} \geq \theta \mathbf{p}_t$ and $\mathbf{p}_t^{2,2} \geq \theta \mathbf{p}_t$. This concludes the necessity part.

Sufficiency. To show the reverse, we start from the condition that both data sets S^1 and S^2 must be consistent with the GARP. From Varian (1982), we know that consistency of $S^1 = \{(\mathbf{p}_t^{1,1}, \mathbf{p}_t^{1,2}; \mathbf{q}_t^1, \mathbf{q}_t^2); t = 1, ..., T\}$ and $S^2 = \{(\mathbf{p}_t^{2,2}, \mathbf{p}_t^{2,1}; \mathbf{q}_t^1, \mathbf{q}_t^2); t = 1, ..., T\}$ with GARP is equivalent to the existence of utility numbers u_t^m and Lagrange multipliers η_t^m such that for $m, n = 1, 2$:

$$u_s^m - u_t^m \leq \eta_t^m \mathbf{p}_t^{m,m\prime}(\mathbf{q}_s^m - \mathbf{q}_t^m) + \eta_t^m \mathbf{p}_t^{m,n\prime}(\mathbf{q}_s^n - \mathbf{q}_t^n).$$

By using these Afriat-like inequalities, we can construct utility functions U^1 and U^2 that rationalise the observed data. For any pair of quantity vectors $(\mathbf{z}^1, \mathbf{z}^2)$, we can define (for $m = 1, 2$)

$$U^m(\mathbf{z}^1, \mathbf{z}^2) = \min_{s \in \{1,\ldots,T\}} [U_s^m + \eta_s^m[(\mathbf{p}_s^{m,1\prime}\mathbf{z}^1 + \mathbf{p}_s^{m,2\prime}\mathbf{z}^2) - (\mathbf{p}_s^{m,1\prime}\mathbf{q}_s^1 + \mathbf{p}_s^{m,2\prime}\mathbf{q}_s^2)]].$$

Let us show that these utility functions effectively provide a cooperative rationalisation with $\theta-$selfishness. First of all, Varian (1982) has proven that $U^m(\mathbf{q}_t^1, \mathbf{q}_t^2) = U_t^m$. Then, for strictly positive μ_t^m, we can simply add up the utility functions of different group members and obtain the following condition:

$$\begin{aligned}\sum_{m,n=1,2,m\neq n} \mu_t^m U^m(\mathbf{z}^m, \mathbf{z}^n) \quad \leq \quad & \sum_{m,n=1,2,m\neq n} \mu_t^m [U_t^m + \eta_t^m [(\mathbf{p}_t^{m,m\prime} \mathbf{z}^m + \mathbf{p}_t^{m,n\prime} \mathbf{z}^n) \\ & -(\mathbf{p}_t^{m,m\prime} \mathbf{q}_t^m + \mathbf{p}_t^{m,n\prime} \mathbf{q}_t^n)]].\end{aligned}$$

For the remainder of the proof, we set $\mu_t^m = 1/\eta_t^m$ and thus we have

$$\begin{aligned}\sum_{m,n=1,2,m\neq n} \mu_t^m U^m(\mathbf{z}^m, \mathbf{z}^n) \quad \leq \quad & \sum_{m,n=1,2,m\neq n} \mu_t^m U_t^m + [(\mathbf{p}_t^{m,m\prime} \mathbf{z}^m + \mathbf{p}_t^{m,n\prime} \mathbf{z}^n) \\ & -(\mathbf{p}_t^{m,m\prime} \mathbf{q}_t^m + \mathbf{p}_t^{m,n\prime} \mathbf{q}_t^n)].\end{aligned}$$

Take any $(\mathbf{z}^1, \mathbf{z}^2)$ that satisfy $\mathbf{p}_t'\mathbf{z}^1 + \mathbf{p}_t'\mathbf{z}^2 \leq \mathbf{p}_t'\mathbf{q}_t^1 + \mathbf{p}_t'\mathbf{q}_t^2$. Then

$$\begin{aligned} & (\mathbf{p}_t^{1,1\prime}\mathbf{z}^1 + \mathbf{p}_t^{1,2\prime}\mathbf{z}^2) - (\mathbf{p}_t^{1,1\prime}\mathbf{q}_t^1 + \mathbf{p}_t^{1,2\prime}\mathbf{q}_t^2) \\ & +(\mathbf{p}_t^{2,2\prime}\mathbf{z}^2 + \mathbf{p}_t^{2,1\prime}\mathbf{z}^1) - (\mathbf{p}_t^{2,2\prime}\mathbf{q}_t^2 + \mathbf{p}_t^{2,1\prime}\mathbf{q}_t^1) \\ = \quad & \mathbf{p}_t'\mathbf{z}^1 + \mathbf{p}_t'\mathbf{z}^2 - \mathbf{p}_t'\mathbf{q}_t^1 - \mathbf{p}_t'\mathbf{q}_t^2 \\ \leq \quad & 0.\end{aligned}$$

The equality follows from the fact that $\mathbf{p}_t^{1,1} + \mathbf{p}_t^{2,1} = \mathbf{p}_t = \mathbf{p}_t^{1,2} + \mathbf{p}_t^{2,2}$. Using this inequality, we finally obtain

$$\sum_{m,n=1,2,m\neq n} \mu_t^m U^m(\mathbf{z}^m, \mathbf{z}^n) \leq \sum_{m,n=1,2,m\neq n} \mu_t^m U_t^m = \sum_{m,n=1,2,m\neq n} \mu_t^m U(\mathbf{q}_t^m, \mathbf{q}_t^n).$$

This shows that $(\mathbf{q}_t^1, \mathbf{q}_t^2)$ maximises the group's objective function subject to $\mathbf{p}_t'\mathbf{z}^1 + \mathbf{p}_t'\mathbf{z}^2 \leq \mathbf{p}_t'\mathbf{q}_t^1 + \mathbf{p}_t'\mathbf{q}_t^2$. As such we have constructed a pair of utility functions that rationalise the data.

To finish the proof we finally need to show that our constructed utility functions also satisfy

$$\varepsilon \leq \frac{\mu_t^1 \frac{\partial U^1}{\partial z_j^1}}{\mu_t^2 \frac{\partial U^2}{\partial z_j^1}} \text{ and } \varepsilon \leq \frac{\mu_t^2 \frac{\partial U^2}{\partial z_j^2}}{\mu_t^1 \frac{\partial U^1}{\partial z_j^2}} \quad \text{with } j = 1, \ldots, n.$$

In order to do this, we use that (with $m, n = 1, 2, m \neq n$) $\mathbf{p}_t^{m,m} \geq \theta \mathbf{p}_t, \mathbf{p}_t^{n,m} \leq (1-\theta)\mathbf{p}_t$ and $\varepsilon = \dfrac{\theta}{1-\theta}$ which shows that

$$\varepsilon = \frac{\theta}{1-\theta} \leq \frac{p_{t,j}^{m,m}}{p_{t,j}^{n,m}}.$$

Above, we have constructed utility functions $U^m(\mathbf{q}^1, \mathbf{q}^2)$ from the Afriat inequalities. These piecewise linear functions are monotonic and concave. Deriving these utility functions with respect to $\mathbf{q}^1$ and $\mathbf{q}^2$ at $(\mathbf{q}_t^1, \mathbf{q}_t^2)$ gives the following inequalities for the supergradients:

$$\begin{aligned} \frac{\partial U^m}{\partial \mathbf{q}_t^m} &\leq \eta_t^m \mathbf{p}_t^{m,m}, \\ \frac{\partial U^n}{\partial \mathbf{q}_t^m} &\leq \eta_t^n \mathbf{p}_t^{n,m}. \end{aligned}$$

Using $\mu_t^m = 1/\eta_t^m$ and $\varepsilon \leq \dfrac{p_{t,j}^{m,m}}{p_{t,j}^{n,m}}$, $m, n = 1, 2$, and $m \neq n$, we effectively obtain that

$$\varepsilon \leq \frac{p_{t,j}^{m,m}}{p_{t,j}^{n,m}} = \frac{\dfrac{1}{\eta_t^m} \dfrac{\partial U^m}{\partial q_{t,j}^m}}{\dfrac{1}{\eta_t^n} \dfrac{\partial U^n}{\partial q_{t,j}^m}} = \frac{\mu_t^m \dfrac{\partial U^m}{\partial q_{t,j}^m}}{\mu_t^n \dfrac{\partial U^n}{\partial q_{t,j}^m}} = \frac{\dfrac{\mu_t^m}{\lambda_t} \dfrac{\partial U^m}{\partial q_{t,j}^m}}{\dfrac{\mu_t^n}{\lambda_t} \dfrac{\partial U^n}{\partial q_{t,j}^m}}.$$

3.B Distribution of individual selfishness for noisy data

Figure 3.5 presents the distribution of the individual selfishness parameter for different degrees of noise/randomness in the data. Specifically, we have replaced the actual choice (and allocation) of a dyad in a particular observation with a random choice from a uniform distribution on the corresponding choice set (as well as a random allocation pattern), with a probability equal to π. Obviously, $\pi = 0$ corresponds to the original distribution presented in Figure 3.1 since no noise is added. On the other hand, $\pi = 1$ would correspond to the distribution of selfishness under Bronars' simulation procedure. The results (based on 500 simulations per value of π) suggest that the distribution of the individual selfishness parameter is relatively robust up to and including $\pi = 0.2$. This indicates that our results are robust to limited degrees of randomness.

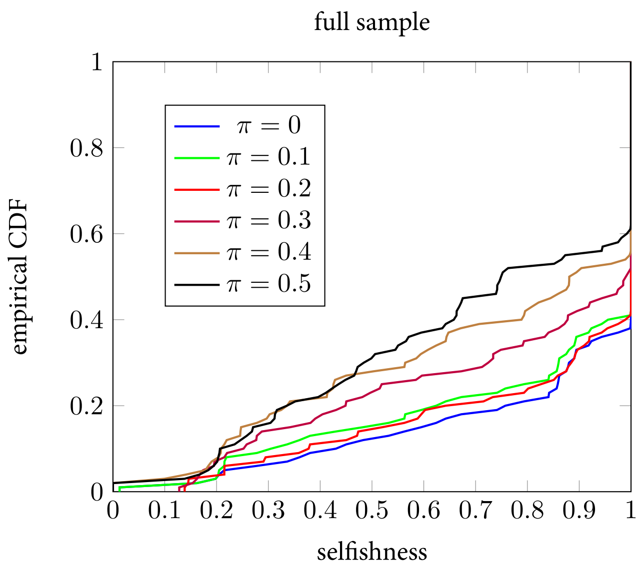


Figure 3.5: Cumulative distribution of the individual selfishness parameter with noise; whole sample with probability π that a choice is replaced with a random bundle

3.C Kolmogorov-Smirnov tests

As a robustness check for our ranksum tests, we also conduct Kolmogorov-Smirnov tests. The results are reported in Table 3.8. For each child characteristic (age, gender and friendship) we carry out three different tests (which in Table 3.8 correspond to the three rows for each child characteristic). The first test uses the alternative hypothesis that θ^m values are smaller in group 0 than in group 1. The second uses the (opposite) alternative hypothesis that θ^m is systematically smaller in group 1 than in group 0. The final test uses the non-directional alternative hypothesis.

Interestingly, the results in Table 3.8 generally confirm the Wilcoxon test results that we presented in the main text. First, our Kolmogorov-Smirnov tests again indicate that the distribution of the θ^m-parameter depends on age. In particular, the null hypothesis that there is no effect is rejected in favor of the alternative hypothesis that sixth graders are more selfish than younger children. Next, we find a weak effect of friendship between dyad members. Specifically, the null hypothesis is rejected in favor of the alternative hypothesis that children who are strong friends with their partner will be less selfish in their consumption behaviour. Finally, and similar to before, the effect of gender is not really significant.

Kolmogorov-Smirnov	Smaller group	D	P-value	Corrected
age	0 kind.,thirdg.	0.3084	0.014	
	1 sixthg.	0.0000	1.000	
	Combined	0.3084	0.028	0.016
friendship	0 weak	0.0000	1.000	
	1 strong	-0.2210	0.108	
	Combined	0.2210	0.215	0.157
gender	0 girls	0.1767	0.217	
	1 boys	0.0000	1.000	
	Combined	0.1767	0.429	0.348

Table 3.8: Kolmogorov-Smirnov tests

Chapter 4

Revealed preferences for diamond goods[1]

4.1 Introduction

Economics researchers are typically confronted with a tradeoff between the (psychological) realism of their models and the ability of the models to describe and estimate the underlying structure of consumers' decisions (Rabin, 2013). For instance, it is well known that consumers care not only about the quantity of their purchases but also for the value associated with various commodities (Ng, 1987; Mandel, 2009). However, allowing for price-dependent preferences generally weakens the testable implications of a model (unless additional assumptions are made on the functional form of utility functions, for instance). Here we present a model in which consumers are allowed to care about the value of a purchase (diamond effect). We present a corresponding revealed preference characterisation, which enables us to test rationality in the presence of diamond effects. We also present a first (non-parametric) empirical application of our model with diamond effects.

[1] I refer to the working paper version of Cosaert (2013).

Diamond goods and price-dependent preferences Demand analysis typically treats consumers as 'rational' or 'optimising' agents who maximise their utility by purchasing the commodities they like. The rationality assumption enables researchers to estimate welfare-related measures (such as cost-of-living indices) and demand functions for various commodities on the basis of real expenditure data. The assumption that consumers 'maximise' their utility is therefore crucial. In most applications, the utility functions are rather strictly defined: it is assumed that consumers care about the quantity of their purchase. However, Ng (1987, 1993) argues that sometimes a good is purchased for its value rather than for its intrinsic consumption effect. Jewellery is probably the most intuitive example. A diamond is not always purchased for its size. Its value may be more important as a means to please a loved one. Similarly, an art collection is prised for its value rather than the number of pieces in the collection (Mandel, 2009). More generally, various goods can have some degree of this so-called diamondness. When individuals treat their friends to dinner, go shopping in expensive clothing stores, or acquire a collection of wines or cigars, we can reasonably argue that these people care about the value of their purchase. Testing the nature of commodities to see whether they can be described as diamond goods is important. Ng (1987) shows that the rules for optimal taxation of diamond goods are very different from the rules for optimal taxation of standard goods. From a theoretical perspective, taxing these goods increases government revenues without imposing an overly large burden on consumers. Moreover, failure to take diamond effects into account can lead to biased results for rationality, welfare, and demand (Heffetz and Shayo, 2009).

Preferences for value can have different sources. Consumers may have preferences for value because they believe that value signals quality (the quality effect), because they want to portray their wealth by purchasing expensive items (the status effect or conspicuous consumption) or because they have an internal desire to possess expensive items (the diamond effect in the narrow sense, following Ng (1987, 1993)). The distinction between diamond and status effects is rather formal. Prices affect utility either directly (because consumers care about the value of the commodities they possess) or indirectly (when consumers derive utility from value because value is observable to society). In the latter case, visibility of the expenditure plays an important role. In this study, we focus on the identification of preferences for value for different commodities. We do not formally disentangle the different sources of preferences for value. However, given the specific set-up of the application, we expect that status effects are important. In the application we use aggregated commodity groups. While it is hard to imagine that a price change in a whole category of goods affects a consumer's quality judgements, it is not unthinkable that the consumer's utility responds differently to price changes in more visible (vis-à-vis less visible) product groups provided that he or she cares about society's perception. We link our diamondness results to a visibility index proposed by Heffetz (2011) in order to investigate this argument.

To date, there has been little research that both incorporates price-dependent preferences in a standard model of consumer behaviour and tests the modified model on the basis of observational data. One reason is that it is difficult to disentangle non-budget-constraint from budget-constraint price effects. Moreover, the way in which non-budget-constraint prices impact on decisions is not observed. Heffetz and Shayo (2009), on the one hand, deal with this issue by setting up an experiment in which distinct prices are presented to respondents: relative prices that monitor the choice set and visual price stickers that capture non-budget-constraint price effects. However, in observational data sets, which are typically used for demand and welfare estimation, this type of information is unavailable. Basmann et al. (1988), on the other hand, try to elicit Veblen effects (measured as elasticity of the marginal rate of substitution with respect to total expenditures) from observational data by estimating a Fechner-Thurstone direct utility function. This utility function has both quantity and price as arguments, so price-dependent preferences can be incorporated. However, much structure is imposed on the utility functions. This has two potential drawbacks. First, if the method rejects rationality, it is uncertain whether the individual was truly irrational or whether an incorrect specification of the utility function was imposed. Second, to estimate the utility functions it is assumed that consumers are homogeneous in terms of preferences. To deal with these problems, we follow a nonparametric (revealed preference) approach.

Revealed preference To obtain a test of rationality without having to specify individual utility functions, we use the revealed preference approach. Revealed preference models in the tradition of Samuelson (1938), Afriat (1967), Diewert (1973), and Varian (1982) define refutable conditions that need to hold in order for a consumer to be rational. The conditions are derived from a finite set of observables: price and quantity information at different points in time. Under the assumption of preference homogeneity over time, revealed preference theory allows testing of the transitivity of preference relations without imposing a functional form on the utility functions. Another attractive feature of this methodology is that consumption decisions of different agents can be analysed independently, thereby fully recognising that different agents can have different tastes.

Here we argue that letting preferences depend on both quantity and expenditure does not automatically preclude a revealed preference test. However, we need to modify the test. Standard revealed preference conditions are unable to take diamond effects into account. It is typically assumed that individuals care about quantity (or at least the intrinsic characteristics of the good[2]) but not about the total value of a purchase. Failure to model additional price effects can lead to incorrect conclusions on the rationality of consumers. Consumer choices that seem irrational according to the standard test may be rationalisable if diamond effects are taken into account, and vice versa.

Unfortunately, the theory of revealed preference and the conjecture that preferences depend on value or price are difficult to reconcile (Bilancini, 2011; Frank and Nagler, 2012). Revealed preference theory requires a finite data set of consumption choices under different price regimes while maintaining a constant preference ordering. If the preference ordering itself is influenced by prices, revealed preference theory becomes useless because we cannot

[2]Blow et al. (2008) provide a revealed preference analysis of characteristics models.

compare different consumption bundles over time.[3]

For this reason we focus on preferences for value. The preferences that we consider here are a special case of price-dependent preferences[4]. Prices enter as a factor of quantities in a homogeneous utility function. Homogeneity of the utility function implies that welfare comparisons (over time) are possible, hence that the revealed preference approach is valid. Moreover, using value as an argument of the utility function imposes meaningful restrictions on the observed choices. After all, the marginal utility from value is related to the consumed quantity of the corresponding commodity via first-order conditions. This allows us to develop testable revealed preference conditions that take price-dependent preferences into account, at the expense of losing some generality. These tests are useful to identify the degree of diamondness associated with commodities.

Contribution This paper makes theoretical, methodological, and empirical contributions to the existing literature.

First, at a theoretical level we present a model in which consumers are allowed to care about both the quantity and value associated with commodities. Our model builds on the framework developed by Ng (1993) but generalises by allowing for more than one 'diamond' good. Moreover, we introduce a parameter that captures the diamondness of commodities,

[3]Pollak (1977) provides an insightful overview of the modelling of price-dependent preferences. The author considers two distinct ways in which price-dependent preferences are analysed: the unconditional approach and the conditional approach. According to the unconditional approach, economic agents express their preferences not only over quantities but also over price–quantity pairs (Kalman, 1968; Piccione and Rubinstein, 2008). In this case, welfare conclusions are still possible (there exists a homogeneous preference ordering defined over quantities and prices) but the model cannot be tested on the basis of data from standard consumption surveys, in which individuals choose quantities and not price–quantity pairs. Because data on choices over price–quantity pairs (at any moment in time) are generally unavailable, the conditional approach seems more popular in empirical work. Following this approach, the preference orderings (defined over quantities) are conditional on prices. A popular functional specification for the utility function — which accounts for preference-shifting parameters such as price — is the generalised Fechner-Thurstone utility function. However, this utility function does not allow for welfare comparisons between periods in which preference-shifting parameters (i.e. prices) take different values.

[4]Note that even more generally, prices can also impact on consumption decisions beyond their effects through budget constraints *or* preferences. Chiappori (1988, 1992) and Apps and Rees (1988) developed a collective model of consumption to describe consumption decisions by households. The collective model is theoretically attractive for studying joint decision-making because it allows different members to have different preferences and lets intra-group bargaining power vary over time (the collective model and its revealed preference characterisation is discussed in more detail in Chapter 3). Interestingly, variation in bargaining power may be driven by variation in the market prices of commodities.

which is the marginal willingness to pay for value. Capturing the diamondness of a commodity in a single parameter is in accordance with research by Rabin (2013), who supports a PEEM (portable extension of existing models) approach, whereby one parameter is added to existing economic models to incorporate insights from the behavioural literature. By letting the diamondness of a commodity vary between 0 and 1, we can move from a model with traditional preferences (consumers only care about quantity) to the model of Ng (1987), which treats a commodity as a *pure* diamond good (consumers only care about the value associated with the *pure* diamond good). If the diamondness lies strictly between 0 and 1, the marginal willingness to pay for additional units of consumption stems from both the intrinsic utility associated with the good and its value.

Second, at a methodological level we present the corresponding revealed preference characterisation. The papers by Ng (1987, 1993) are purely theoretical. The author described a model according to which consumers care about the value of one good and derived rules for optimal taxation (of the diamond good). In this chapter, we also focus on the implementation and testing of a model with diamond goods. We show that, conditional on some level of diamondness, rationality can still be tested in a meaningful way. Moreover, the different characterisations are generally non-nested. In other words, increasing diamondness associated with one or multiple goods does not automatically relax the revealed preference conditions. This implies that we can construct 'bounds' on the diamondness associated with various commodities.

Finally, at an empirical level we apply our revealed preference tests for rationality to a data sample from the Russian Longitudinal Monitoring Survey (RLMS). To the best of our knowledge, this is the first empirical revealed preference test of diamond effects.

In Section 4.2, we define rationality and preferences for value and introduce our *diamondness* parameter. In Section 4.3, we present the revealed preference characterisations that allow us to test rationality for different specifications of the diamondness vector. These characterisations contain conditions that can be implemented using (mixed integer) linear programming techniques. We also discuss the non-nestedness of our characterisations. This section ends with a discussion of standard measures of empirical performance in the revealed preference literature. In Section 4.4, we briefly discuss our sample taken from the RLMS. Section 4.5 presents rationality results under different specifications of the diamondness vector. We also show that the appropriate degree of diamondness depends on the individual and the product at hand. Section 4.6 concludes.

4.2 Theory

We present a model in which consumers derive utility from both the goods they consume and the expenditure on these commodities. We follow a so called unconditional approach, in which preferences are homogeneous over time (i.e. the preference parameters are independent of prices) although prices are allowed to enter the utility function. We let prices enter the utility function in a specific way, namely as a factor of quantities. Then the marginal utility from value is directly related to the observed quantities via first-order conditions. This makes that the revealed preference approach is valid and that it provides us with refutable conditions.

Ng (1987) introduced a model in which one good is assumed to be a (pure) diamond good, that is, the market value of this commodity enters the consumer utility function. In a later paper, Ng (1993) also allowed for mixed diamond goods, whereby both the market value and the intrinsic consumption component of one (mixed) diamond good enter the utility function. Unfortunately, in the framework of Ng (1987, 1993) there is only one pure or mixed diamond good, and other goods are standard goods. Here we present a framework in which, in principle, *all* goods can be characterised by some degree of diamondness, and diamondness (expressed on a continuum between 0 and 1) can be measured in monetary terms.

4.2.1 Modelling diamondness

To construct a testable model with diamond effects, we build on the model of Ng (1993), which is itself a generalisation of the original Ng (1987) framework. The model of Ng (1993) assumes that utility can be derived from both the quantity and market price value of one commodity. We extend this framework to assess preferences for value associated with multiple goods, and we introduce a parameter to capture the relative importance of the marginal willingness to pay for value.

Suppose that we have a data set $S = \{\mathbf{P}_t, \mathbf{Q}_t \,|\forall t \in T\}$ consisting of $|T|$ observations. For each observation t, this data set contains information on the observed quantity vector $\mathbf{Q}_t \in \mathbb{R}_{+}^{|N|}$ as chosen by consumers and the corresponding price vector $\mathbf{P}_t \in \mathbb{R}_{++}^{|N|}$. Let $\mathbf{M}_t \in \mathbb{R}_{+}^{|N|}$ represent the vector of expenditures on $|N|$ commodities in period t (i.e. $\mathbf{M}_t$ consists of elements $M_t^n = P_t^n Q_t^n$). We assume that utility functions take the form $U(\mathbf{Q}, \mathbf{M})$, which means that consumers can derive utility from quantity on the one hand and from the value of a purchase on the other hand:

$$\max_{\mathbf{Q}} U(\mathbf{Q}, \mathbf{M})$$

$$s.t.$$

$$\lambda_t : \mathbf{P}_t' \mathbf{Q} \leq y_t$$

$$\forall n \in N : M^n = P_t^n Q^n.$$

We obtain testable implications by deriving the first-order conditions associated with the above problem:

$$\forall n \in N \ : \frac{\partial U(\mathbf{Q}_t, \mathbf{M}_t)}{\partial Q^n} + \frac{\partial U(\mathbf{Q}_t, \mathbf{M}_t)}{\partial M^n} P_t^n = \lambda_t P_t^n. \tag{4.1}$$

This expression clearly shows that the total marginal utility from an additional unit of good n is composed of the marginal utility associated with the good itself, $\frac{\partial U(\mathbf{Q}_t,\mathbf{M}_t)}{\partial Q^n}$, and the marginal utility associated with the value of this good, $\frac{\partial U(\mathbf{Q}_t,\mathbf{M}_t)}{\partial M^n}$. The second term follows directly from the *diamond effect*. We define diamondness as the *marginal willingness to pay* for value. To this end, we divide the marginal utility from additional expenditure, $\frac{\partial U(\mathbf{Q}_t,\mathbf{M}_t)}{\partial M^n}$, by the marginal utility from one unit of income, λ_t. In this way, we obtain a measure for the diamond effect in monetary terms:

$$\theta^n = \frac{1}{\lambda_t}\frac{\partial U(\mathbf{Q}_t,\mathbf{M}_t)}{\partial M^n}.$$

We assume that diamondness θ^n is time-invariant. Although this assumption is not necessary for the derivation of revealed preference conditions, it is convenient because it allows us to identify one diamondness parameter per commodity. This avoids a grid search on multiple diamondness parameters for the same commodity. Moreover, even with fixed diamondness, the model generalises the standard utility maximisation framework (note in this respect that the standard model has diamondness fixed equal to 0) without being overly permissive. One way to think about this assumption is to assert that diamondness is specific to a commodity, irrespective of the quantity consumed.[5]

Interestingly, we let the diamondness parameter θ^n be commodity-specific. We allow for the fact that certain commodities are more likely to trigger diamond effects than others. The diamondness parameter is bounded between 0 and 1. This stems from the assumptions that neither the marginal utility from quantity nor the marginal utility from value can be negative. These assumptions, in combination with Condition (4.1) and the definition of the

[5] At this point, it is worth noting that θ is not identified for the generic Cobb Douglas function. In other words, the outcome of the maximisation of a generic Cobb Douglas utility function is independent of θ. For this reason, different diamondness values are empirically indistinguishable in the Cobb Douglas framework. The reason is that the maximisation of the Cobb Douglas utility function on the one hand and the constant θ assumption on the other hand both imply that the expenditures are proportional to income. For more general specifications - such as Constant Elasticity of Substitution - this non-identification result no longer holds. In particular, the maximisation of a generic CES utility function implies that expenditures are a function of preference parameters, income *and* prices, while the constant θ assumption implies that expenditures are a function of preference parameters and income alone.

diamondness parameter, automatically imply that $\theta^n \in [0,1]$.

Suppose first that $\theta^n = 1$, in other words, the marginal utility from an additional unit of income stems purely from the diamond effect. In this case, good n is a *pure diamond* good (Ng, 1987). Consumers derive utility only from expenditure on this good and not from its quantity. Second, suppose that $\theta^n = 0$, in other words, additional expenditure on good n has no direct impact on the utility of the consumer. In this case, good n is a *standard* good, in the sense that additional income only impacts utility because the consumer can purchase larger amounts of good n. Finally, we allow the diamondness weight θ^n to take any value between 0 and 1, $\theta^n \in [0,1]$, which accounts for the possibility that goods are purchased *both* for their intrinsic value *and* for their monetary value.

Interestingly, the proposed model encompasses the neoclassical consumption model when $\theta^n = 0$ for all $n \in N$. It can easily be verified that the second term in Condition (4.1) (i.e. the diamond effect) would drop out for all commodities. The proposed model therefore fits within the PEEM framework of Rabin (2013). Indeed, our model extends the neoclassical counterpart only insofar as the newly proposed diamondness vector deviates from 0. At the other extreme, our model encompasses the situation in which all commodities are purchased only for their value, when $\theta^n = 1$ for all $n \in N$. In this case, it is always possible to construct a (consistent) preference ordering such that utility increases with total expenditure.

Assume now that $\theta \in \mathbb{R}^{|N|}_{[0,1]}$ is a vector containing the $|N|$ diamondness parameters. Utility-maximising behaviour conditional on the diamondness vector is described in Problem 4.1.

Problem 4.1. Optimisation problem $OPT - \theta$:

$$\max_{\mathbf{Q}} U(\mathbf{Q}, \mathbf{M})$$

$$s.t.$$

$$\lambda_t : \mathbf{P}_t'\mathbf{Q} \leq y_t$$

$$\forall n \in N : \left\{ \begin{array}{c} M^n = P_t^n Q^n \\ \theta^n = \frac{1}{\lambda_t} \frac{\partial U(\mathbf{Q}_t, \mathbf{M}_t)}{\partial M^n}. \end{array} \right\}$$

Revealed preference theory then proceeds by looking for some utility function U such that the observed consumption pattern solves Problem 4.1. Definition 4.2 links rationalisability with $\theta - diamondness$ to consistency with the above program.

Definition 4.2. Consider a data set $S = \{\mathbf{P}_t, \mathbf{Q}_t \, | \forall t \in T\}$. We say that S is rationalisable with $\theta - diamondness$ if there exists a utility function U defined over quantity vector $\mathbf{Q}$ and expenditure vector $\mathbf{M}$ such that $\{\mathbf{Q}_t; \forall t \in T\}$ solves optimisation Problem 4.1.

Before presenting a nonparametric methodology to test consistency with Definition 4.2, let us briefly discuss the implications of the introduction of diamond effects for the Slutsky matrix. Towards this end, suppose that we observe, or can estimate, the uncompensated demand for all goods conditional on θ. Let $f^n(\mathbf{P}, y, \theta)$ be the uncompensated demand for good n, conditional on price vector $\mathbf{P}$, income level y and diamondness vector θ. Note furthermore that θ is constant and independent of prices $\mathbf{P}$ or expenditures y in the current study[6].

Then the Envelope Theorem implies that

[6]When diamondness depends on prices and/or expenditures, $\theta^n(\mathbf{P}, y)$, the derivation of the Slutsky matrix is in line with the corresponding derivation for a collective model of consumption (see Browning and Chiappori (1998)). However, the condition will be more intricate than the so called SR1 condition for collective rationality because 1) a price change might impact on the diamondness of N commodities and 2) the derivative of the expenditure function no longer corresponds to the Hicksian demand. The latter implies that only a transformation of the matrix containing (price and income) derivatives of the uncompensated demand functions is symmetric and negative semi-definite.

$$\frac{\partial E(\mathbf{P},u,\theta)}{\partial P^i} = (1-\theta^i)h^i(\mathbf{P},u,\theta). \tag{4.2}$$

in which $E(\mathbf{P},u,\theta)$ represents the expenditure function conditional on price vector $\mathbf{P}$, utility level u and diamondness vector θ. The compensated (Hicksian) demand for good n is given by $h^n(\mathbf{P},u,\theta)$. Condition (4.2) deviates from the classical result that $\frac{\partial E(\mathbf{P},u)}{\partial P^i} = h^i(\mathbf{P},u)$. After all, the increase in expenditures necessary to obtain the same level of utility after a price increase is offset by the fact that higher prices increase utility (if $\theta^i > 0$). It can be shown that

$$\frac{\partial^2 E(\mathbf{P},u,\theta)}{\partial P^i \partial P^j} = (1-\theta^i)\frac{\partial h^i(\mathbf{P},u,\theta)}{\partial P^j}$$
$$= (1-\theta^i)\left\{\frac{\partial f^i(\mathbf{P},E(\mathbf{P},u,\theta),\theta)}{\partial P^j} + (1-\theta^j)\frac{\partial f^i(\mathbf{P},E(\mathbf{P},u,\theta),\theta)}{\partial y}f^j(\mathbf{P},E(\mathbf{P},u,\theta),\theta)\right\}.$$

Finally one can see that the uncompensated demands f stem from a rational consumer if and only if the matrix Δ is negative semi-definite and symmetric, with

$$\Delta_{ij} = (1-\theta^i)\Gamma_{ij}$$
$$\Gamma_{ij} = \frac{\partial f^i}{\partial P^j} + (1-\theta^j)\frac{\partial f^i}{\partial y}f^j$$

Notice that these conditions coincide with the conventional Slutsky conditions when $\theta^n = 0$ for all goods n. Testing these conditions requires the estimation of demand functions f which pool the observed demand of different consumers. Our nonparametric test, which we present in the next section, requires only a finite set of consumption choices and avoids the pooling of different, potentially heterogeneous, consumers.

4.3 Methodology

In the previous section we formulated a generalisation of the original models by Ng (1987, 1993). We also introduced a parameter to capture the diamond effect. In this section, we first show how nonparametric (revealed preference) conditions can be derived to test the rationalisability concepts. The revealed preference approach is particularly attractive in the current setting. It avoids putting additional structure on the utility function $U(\cdot)$. This guarantees that recovery of the diamondness vector is independent of the functional form of a utility function. Moreover, it rules out issues related to unobserved heterogeneity across consumers, because each consumer is analysed separately. Next, we show that different specifications of the diamondness vector are generally non-nested. Finally, we discuss standard revealed preference measures that allow us to test the empirical performance of our proposed method.

Revealed preference methodology To set the stage, we first present the revealed preference test for consistency with a standard utility function of the form $U(\mathbf{Q})$. This standard revealed preference test was developed by Afriat (1967), Diewert (1973) and Varian (1982). Specifically, the data set $S = \{\mathbf{P}_t; \mathbf{Q}_t \,|\forall t \in T\}$ is said to be rationalisable if there exists a utility function $U(\mathbf{Q})$ such that

$$\mathbf{Q}_t = \arg\max_{\mathbf{Q}} U(\mathbf{Q}) \text{ s.t. } \mathbf{P}_t'\mathbf{Q} \leq \mathbf{P}_t'\mathbf{Q}_t$$

No further functional form restrictions are imposed on $U(\cdot)$. Afriat's Theorem, a central result in the revealed preference literature, then provides a rationality test.

Proposition 4.3. Consider a data set $S = \{\mathbf{P}_t; \mathbf{Q}_t \,|\forall t \in T\}$. The following conditions are equivalent:

1. The data set S is rationalisable.

2. For all decision situations $t \in T$ and for all commodities $n \in N$, there exist utility numbers u_t, and (Lagrange) multipliers $\lambda_t \in \mathbb{R}_{++}$ such that for all $t, v \in T$:

$$u_t - u_v \leq \lambda_v \mathbf{P}'_v(\mathbf{Q}_t - \mathbf{Q}_v)$$

3. For all decision situations $t \in T$ and for all commodities $n \in N$:

$$S = \{\mathbf{P}_t; \mathbf{Q}_t | \forall t \in T\} \text{ satisfies the } GARP$$

Statement 2 contains the so called Afriat inequalities. Observed behaviour can be rationalised by the standard utility maximisation model if and only if the data are consistent with the Afriat inequalities. Moreover, these conditions are also equivalent to stating that $S = \{\mathbf{P}_t; \mathbf{Q}_t \, | \forall t \in T\}$ is consistent with the Generalised Axiom of Revealed Preference (GARP).

In this section, we propose modified revealed preference conditions that can be used to test consistency with Definition 4.2.

4.3.1 Revealing preferences for diamond goods

We develop a test for rationalisability for a given vector of diamondness weights θ. We investigate whether it is possible to construct a well-behaved utility function U, with both values $\mathbf{M}$ and quantities $\mathbf{Q}$ as arguments, such that the observed consumption pattern $\{\mathbf{Q}_t; \forall t \in T\}$ solves Problem 4.1. We start from the concavity of the utility function. For all t and v, we must have that

$$\begin{aligned} U(\mathbf{Q}_t, \mathbf{M}_t) - U(\mathbf{Q}_v, \mathbf{M}_v) &\leq \sum_{n=1}^{|N|} \frac{\partial U(\mathbf{Q}_v, \mathbf{M}_v)}{\partial Q^n}(Q_t^n - Q_v^n) \\ &\quad + \sum_{n=1}^{|N|} \frac{\partial U(\mathbf{Q}_v, \mathbf{M}_v)}{\partial M^n}(P_t^n Q_t^n - P_v^n Q_v^n). \end{aligned} \tag{4.3}$$

The first-order conditions without restrictions on the diamondness vector were presented in Condition (4.1). They imply that

$$\lambda_t P_t^n = \frac{\partial U(\mathbf{Q}_t, \mathbf{M}_t)}{\partial Q^n} + \frac{\partial U(\mathbf{Q}_t, \mathbf{M}_t)}{\partial M^n} P_t^n. \tag{4.4}$$

We can formulate the marginal utilities in monetary terms by dividing both terms by λ_t. This gives shadow prices $\mathfrak{p}_t^n$ and $\mathfrak{P}_t^n$ such that

$$\begin{aligned} P_t^n &= \mathfrak{p}_t^n + \mathfrak{P}_t^n \cdot P_t^n \\ \mathfrak{p}_t^n &= \frac{1}{\lambda_t} \frac{\partial U(\mathbf{Q}, \mathbf{M})}{\partial Q_t^n} \\ \text{and } \mathfrak{P}_t^n &= \frac{1}{\lambda_t} \frac{\partial U(\mathbf{Q}_t, \mathbf{M}_t)}{\partial M^n}. \end{aligned}$$

We also set $u_t = U(\mathbf{Q}_t, \mathbf{M}_t)$ and $u_v = U(\mathbf{Q}_v, \mathbf{M}_v)$. Finally, our definition of θ^n implies that $\mathfrak{P}_t^n = \theta^n$. By combining the above restrictions, we obtain the necessary conditions for rationalisability with $\theta - diamondness$. In Appendix 4.A, we show that these conditions are also sufficient. In fact, the concavity conditions (4.3) ensure that the first-order conditions (4.4) are both necessary and sufficient for optimality. We can now formulate Proposition 4.4.

Proposition 4.4. Consider a data set $S = \{\mathbf{P}_t; \mathbf{Q}_t \mid \forall t \in T\}$. The following conditions are equivalent:

1. The data set S is rationalisable with $\theta - diamondness$.

2. For all decision situations $t \in T$ and all commodities $n \in N$, there exist utility numbers u_t, (Lagrange) multipliers $\lambda_t \in \mathbb{R}_{++}$, and shadow prices $\mathfrak{p}_t \in \mathbb{R}_{+}^{|N|}$ and $\mathfrak{P}_t \in \mathbb{R}_{+}^{|N|}$ such that for all $t, v \in T$,

$$u_t - u_v \leq \lambda_v \sum_{n=1}^{|N|} \mathfrak{p}_v^n \cdot (Q_t^n - Q_v^n) + \lambda_v \sum_{n=1}^{|N|} \mathfrak{P}_v^n \cdot (P_t^n Q_t^n - P_v^n Q_v^n)$$

$$\forall n \in N : \left\{ \begin{matrix} P_v^n = \mathfrak{p}_v^n + \mathfrak{P}_v^n \cdot P_v^n \\ \mathfrak{P}_v^n = \theta^n. \end{matrix} \right\}$$

3. For all decision situations $t \in T$ and for all commodities $n \in N$, there exist shadow prices $\mathfrak{p}_t \in \mathbb{R}_{+}^{|N|}$ and $\mathfrak{P}_t \in \mathbb{R}_{+}^{|N|}$ such that for all $t, v \in T$,

$$S = \{\mathfrak{p}_t, \mathfrak{P}_t; \mathbf{Q}_t, \mathbf{M}_t | \forall t \in T\} \text{ satisfies the } GARP$$

$$\forall n \in N : \left\{ \begin{matrix} P_v^n = \mathfrak{p}_v^n + \mathfrak{P}_v^n \cdot P_v^n \\ \mathfrak{P}_v^n = \theta^n. \end{matrix} \right\}$$

Statement 1 gives the definition of rationality when the magnitude of the diamond effect is given by θ. Statement 2 presents inequalities that allow us to test the presumption of rationalisability with $\theta - diamondness$. The conditions are similar in nature to the well-known Afriat inequalities. However, there are two main differences. First, prices are not observed. We use shadow prices $\mathfrak{p}_v^n$ and $\mathfrak{P}_v^n$ to capture the marginal willingness to pay for quantity and value, respectively. Second, there is an additional set of $|T| \cdot |N|$ conditions, which state that the sum of the marginal willingness to pay for an additional unit of some good and the marginal willingness to pay for value associated with this good (multiplied by the market price) should equal the respective market price. Statement 3 provides us with an alternative test of rationalisability based on the GARP conditions.

Statements 2 and 3 are easily implementable. Notice first that both $\mathfrak{p}_v^n$ and $\mathfrak{P}_v^n$ are determined as soon as θ^n is specified. Conditional on the vector θ, the conditions in Statement 2 are linear in u_t, u_v and λ_v. Therefore, we can use simple linear programming techniques to implement this test. Statement 3 contains the alternative GARP formulation, which is particularly convenient. When θ^n is specified, on one hand, Statement 3 simply requires verification of a set of combinatorial restrictions. It is no longer necessary to formulate (and solve) a programming problem. On the other hand, Statement 3 enables us to check rationalisability even if θ is unspecified before the analysis. This enables us to endogenise the diamondness parameter θ^n (as well as shadow prices $\mathfrak{p}_v^n$ and $\mathfrak{P}_v^n$), and consequently provides bounds on the feasible set of diamondness values. A GARP-based test with unknown diamondness vector can be implemented using a linear programming problem with integer variables.

As a side note, we point out that this framework is also useful for analysis of bad commodities. The standard neoclassical model stipulates that consumers should not spend their budget on a bad commodity, which reduces their (intrinsic) utility. However, when preferences depend on value, rational consumers can purchase additional units of a bad commodity, as long as the marginal utility from its value exceeds the (negative) intrinsic marginal utility. Testing whether a commodity n is bad is now easy. We can simply modify the requirement

$$\mathfrak{p}_v^n < 0$$

in the revealed preference characterisation in Proposition 4.4. The (negative) marginal utility due to quantity is then offset by the (positive) marginal utility from value if $\mathfrak{P}_t^n > 1$. In order to investigate bads, one could allow that the marginal utility from quantity is negative such that θ^n is no longer bounded from above (by 1).

4.3.2 Independence

One of our main contributions is a rationality test when consumers have preferences for value. As argued above, we can test for rationality conditional on a diamondness vector θ, *or* recover the set of diamondness weights that allows rationalisation of (the largest fraction of) the data. At this point, it is worth noting that the different rationality tests (corresponding to different specifications of diamondness) are generally non-nested.

Indeed, it is possible that a data set violates the predictions of the classical model while it is rationalisable with strictly positive diamondness for particular commodities. Likewise, a data set can be consistent with the classical model while it is not rationalisable for some strictly positive specification of θ^n. This suggests that we can construct meaningful bounds on the diamondness. There is one exception, however. The specification in which all goods are treated as 'pure' diamond goods cannot reject rationality (i.e. setting all diamondness parameters to $\theta = 1$ can always trivially rationalise the data). An empirical solution to this problem (discriminatory power) is discussed in Section 4.5. Intuitively, discriminatory power captures the strength of a test, that is, its capacity to reject rationality when confronted with random choice behaviour.

To demonstrate the non-nestedness between different specifications, we use two examples. Figures 4.1 and 4.2 provide graphical illustrations of two different data sets.

Consider first a data set with $|T| = 2$ observations, $|N| = 2$ goods, and price vectors $\mathbf{P}_1 = [2, 1]'$ and $\mathbf{P}_2 = [1/2, 1]'$ and quantity vectors $\mathbf{Q}_1 = [3, 2]'$ and $\mathbf{Q}_2 = [2, 3]'$ for the first and second observation, respectively. The left graph in Figure 4.1 presents the corresponding budget constraints in terms of quantities. After all, consumers who have standard preferences only care for consumed quantities. Given this assumption, however, the choices in this first example are clearly irrational. The consumer preferred bundle $\mathbf{Q}_1$ over bundle $\mathbf{Q}_2$ in the first observation while he preferred $\mathbf{Q}_2$ over $\mathbf{Q}_1$ in the second observation. The right graph in Figure 4.1 differs from the left in that good 1 is presented as a (pure) diamond good, i.e. with $\theta_1 = 1$. As such, consumers care for the quantity associated with good 2 and the value associated with good 1. The budget constraints are re-defined accordingly. The budget constraint of observation 1 tilts outward because the quantities of good 1 are multiplied by the high price of good 1 in observation 1. Similarly, the budget constraint of observation 2 tilts inward because the quantities of good 1 are multiplied by the low price of good 1 in observation 2. Given this alternative assumption driven by the diamondness of good 1, the consumer still preferred bundle $\mathbf{Q}_1$ over bundle $\mathbf{Q}_2$, but the reverse no longer holds. The choices are perfectly rationalisable when good 1 is a diamond good[7].

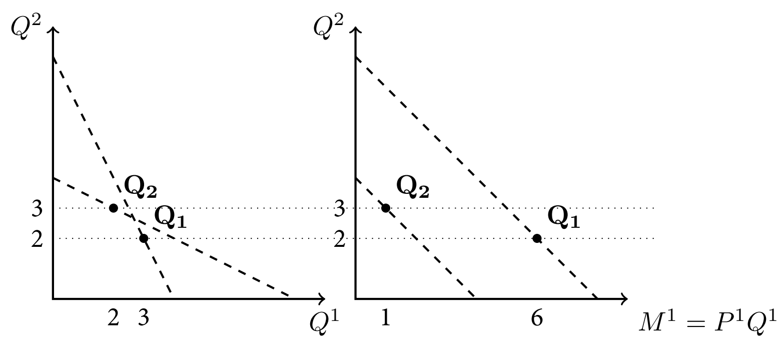


Figure 4.1: Non-nestedness example with 2 goods (left graph: all standard goods, right graph: diamond good 1)

For the reverse, consider a data set with $|T| = 2$ observations, $|N| = 3$ goods, and price vectors $\mathbf{P}_1 = [1, 2, 3]'$ and $\mathbf{P}_2 = [4, 3, 2]'$ and quantity vectors $\mathbf{Q}_1 = [1, 9, 10]'$ and $\mathbf{Q}_2 = [9, 3, 0]'$ for the first and second observation, respectively. Bundle $\mathbf{Q}_1$ was chosen from the budget set bounded by hyperplane $h_1 : y_1 = \mathbf{P}_1'\mathbf{Q}_1$ and $\mathbf{Q}_2$ was chosen from the budget set bounded by hyperplane $h_2 : y_2 = \mathbf{P}_2'\mathbf{Q}_2$. Notice that all bundles on and below the hyperplane associated with a particular observation were, in principle, affordable. The left graph in Figure 4.2 presents the budget constraints in terms of quantities. We can infer from the

[7]Algebraically, one can verify that $[\mathbf{P}_1'\mathbf{Q}_1 = 8] > [\mathbf{P}_1'\mathbf{Q}_2 = 7] \Rightarrow u_1 > u_2$ and $[\mathbf{P}_2'\mathbf{Q}_2 = 4] > [\mathbf{P}_2'\mathbf{Q}_1 = 3.5] \Rightarrow u_2 > u_1$ lead to a contradiction, while the combination of $[\mathbf{P}_1'\mathbf{Q}_1 = 8] > [M_2^1 + P_1^2 Q_2^2 = 3.5] \Rightarrow u_1 > u_2$ and $[\mathbf{P}_2'\mathbf{Q}_2 = 4] < [M_1^1 + P_2^2 Q_1^2 = 8] \not\Rightarrow u_2 > u_1$ is feasible.

left graph that the choices are rational (given the traditional assumption of preferences for quantity). The consumer preferred bundle $\mathbf{Q}_1$ over bundle $\mathbf{Q}_2$ in the first observation ($\mathbf{Q}_2$ was below the budget hyperplane of observation 1 and hence affordable), while the reverse does not hold. The right graph in Figure 4.2 differs from the left in that good 2 is presented as a (pure) diamond good, i.e. with $\theta_2 = 1$. As such, consumers care for the quantity associated with goods 1 and 3 and the value associated with good 2. The budget constraints are re-defined accordingly. Specifically, the budget hyperplanes tilt outward because the prices of good 2 are taken up as arguments of the utility function. Given this alternative assumption driven by the diamondness of good 2, however, we find that the choices are no longer consistent. The consumer preferred bundle $\mathbf{Q}_1$ over bundle $\mathbf{Q}_2$ in the first observation and bundle $\mathbf{Q}_2$ over bundle $\mathbf{Q}_1$ in the second observation ($\mathbf{Q}_1$ from hyperplane h_1 was below hyperplane h_2 and hence affordable). As a result the theory of rational consumption rejects that good 2 is a (pure) diamond good (with $\theta_2 = 1$)[8].

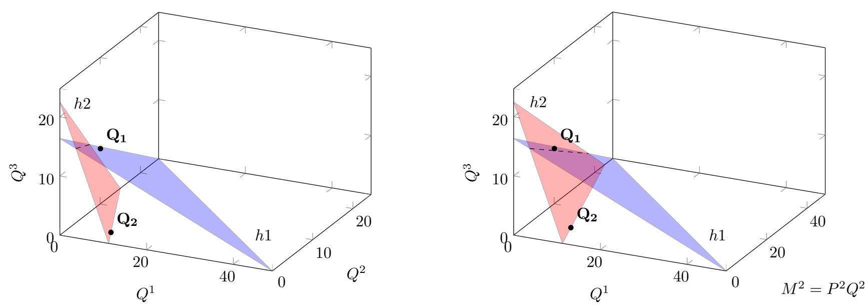


Figure 4.2: Non-nestedness example with 3 goods (left graph: all standard goods, right graph: diamond good 2)

[8] Algebraically, one can verify that the combination of $[\mathbf{P}_1'\mathbf{Q}_1 = 49] > [\mathbf{P}_1'\mathbf{Q}_2 = 15] \Rightarrow u_1 > u_2$ and $[\mathbf{P}_2'\mathbf{Q}_2 = 45] < [\mathbf{P}_2'\mathbf{Q}_1 = 51] \not\Rightarrow u_2 > u_1$ is feasible, while $[\mathbf{P}_1'\mathbf{Q}_1 = 49] > [P_1^1 Q_2^1 + M_2^2 + P_1^3 Q_2^3 = 18] \Rightarrow u_1 > u_2$ and $[\mathbf{P}_2'\mathbf{Q}_2 = 45] > [P_2^1 Q_1^1 + M_1^2 + P_2^3 Q_1^3 = 42] \Rightarrow u_2 > u_1$ lead to a contradiction.

4.3.3 Measures of empirical performance

We conclude this section with an overview of empirical performance measures which will allow us to assess various specifications of the diamondness vector: pass rates, power and predictive success. On the basis of these measures, we will be able to critically examine our revealed preference characterisations when brought to the data.

As indicated earlier, pass rates give the fraction of data sets that are consistent with the corresponding revealed preference conditions.

Given that it is our purpose to compare the empirical performance of specifications which impose various degrees of diamondness, we also need a measure for the strength of our tests. It is easy to show why such measure is necessary. A characterisation where all diamondness parameters are set to 1 would impose no meaningful restrictions. The underlying explanation is that, in such case, a preference ordering $U(\mathbf{Q}, \mathbf{M})$ can be constructed which is increasing with total expenditures $\sum^{n} M^n$ and which trivially rationalises the observed expenditures[9]. Furthermore, even if some goods were modelled as standard goods, the strength of our test could be influenced by the choice of the parameters. In order to control for this, we compute the 'power', d, associated with different specifications of the model (for a review of power measures, see Andreoni et al. (2011)). For robustness, we use two different power measures: the bootstrap power index and Bronars (1987)' power index, respectively. The bootstrap power method simulates random data by drawing budget shares from the distribution of observed budget shares in the sample. Bronars' approach simulates random data by drawing shares from a uniform distribution on the budget hyperplane. However, in order to take the large number of zero expenditures in the data into account, we modify Bronars' approach following Cherchye et al. (2009), who apply Bronars' procedure to a similar sample from the RLMS. Specifically, we impose that each simulated budget share should not exceed the relative number of zero expenditures in the data. In this way,

[9] In this respect, Bilancini (2011) and Frank and Nagler (2012) also noted that *any* pattern of choices can be rationalised by a (non-restricted) utility function of the form $U(\mathbf{Q}, \mathbf{P})$.

we control for commodity groups which are not frequently purchased. Finally, the random shares are multiplied with the consumer's income and divided by the corresponding prices to obtain new (random) commodity bundles. Power is one minus the pass rate of random data sets.

In the end, we want to decide on which specification is most attractive in terms of empirical performance. Therefore, we apply the measure of predictive success, p, proposed by Selten (1991) and discussed in more detail by Beatty and Crawford (2011). An elegant feature of this measure is that it combines pass rates (r) and discriminatory power (d). Predictive success (p) is defined as:

$$p = r - (1 - d)$$

Higher predictive success indicates that the model is better able to distinguish between observed behaviour (which is supposed to be rational according to the model) and random, simulated behaviour (which is supposed to violate the conditions of the model). The more positive predictive success scores are hence desirable.

4.4 Data

We apply our revealed preference tests to consumer data from the RLMS from 1994 to 2006, with the exception of 1997 and 1999. These 11 waves correspond to the second collection phase of the RLMS data (Phase II). We assume that the preferences of each respondent are sufficiently stable over time to construct the revealed preference conditions.

We restrict our attention to data sets for single individuals who do not receive any unemployment benefits. Furthermore, we only consider individuals who report expenditure for the 11 waves. Finally, we focus on individuals who were house and car owners during the full observation period. This yields a sample of 82 individuals. By conditioning on house and car ownership, we can exclude 'large' decisions on durable goods from the analysis.[10] The reason is straightforward: we want to make a clear distinction between decisions driven by the diamond effect on one hand and intertemporal portfolio decisions on the other hand. Since the focus of this paper is on the former, we only consider nondurable commodities.

The nondurable commodities that we consider here are bread, potatoes, vegetables, fruit, meat, dairy products, alcohol, tobacco, food outside the home, clothes, car fuel, wood fuel, gas fuel, and luxury products. This grouping follows Cherchye et al. (2009), who conducted a similar revealed preference application to the RLMS data.

The prices of these aggregates are weighted (geometric) means of the prices associated with various detailed subgroups of goods. For instance, the price of alcohol is a weighted mean of the prices of vodka, liquor, and beer. We therefore have that the aggregate price P_v^n for some commodity group n in period v is equal to

$$P_v^n = \prod^{k} (p_v^k)^{w^k},$$

where index k denotes subgroups of products that belong to the aggregate commodity n, and the weights w^k are determined by the average expenditure share of k relative to expenditure on commodity n. By construction, these weights sum to one.

The aggregation of prices p_v^k in P_v^n (and quantities q_v^k in Q_v^n) is not uncontroversial. Changes in the aggregates may reflect changes in the composition and/or quality of the aggregates. The relationship between the aggregate prices and quantities may therefore stem from the aggregation rather than the diamond effect. Consider for instance an aggregate commodity Q_v^n comprising a good q_v^l and a good q_v^h:

$$Q_v^n = \frac{p_v^l q_v^l + p_v^h q_v^h}{P_v^n} = \frac{p_v^l q_v^l + p_v^h q_v^h}{(p_v^l)^w (p_v^h)^{1-w}}. \tag{4.5}$$

[10] We thus implicitly assume that decisions on these nondurable commodities and large decisions on durables are weakly separable. Although this assumption is contestable, it is quite common in applied static demand analysis. Moreover, interpersonal variation in durable decisions is not an issue because we analyse each agent separately.

From Condition (4.5) we can show that Q_v^n increases with p_v^l if the budget share of l is greater than w, and that Q_v^n increases with p_v^h if the budget share of h is greater than $1-w$.[11] This implies that our results may be sensitive to the aggregation method under consideration. To deal with this issue, we rely on the Hicks Composite Commodity Theorem[12], which was first discussed by Leontief (1936) and Hicks (1946) and further developed by Gorman (1953). The theorem states that commodities Q_v^n and aggregate prices P_v^n can be treated in the same way as goods q_v^l and q_v^h and unobserved prices p_v^l and p_v^h, provided that the relative prices p_v^h/p_v^l remain stable across observations v (i.e., $p_v^h/p_v^l = \alpha$). In Appendix 4.B we show that the Hicks Theorem still holds in a setting with diamond goods. Moreover, our data show strong correlations between the prices of various subgroups that belong to the same aggregate.[13]

For the revealed preference analysis, we restrict our attention to *real* prices. We divide all nominal prices by the average price level in each period. In this way, our approach is consistent with the relative price hypothesis, which postulates that preferences are independent of the nominal units in which prices are measured (see Pollak (1977)). The relative price hypothesis avoids that homogeneous changes in prices over time (changes in the price of all goods, e.g. due to inflation) impact on the diamondness results.

To limit the number of parameters to be estimated in the empirical application, we let diamondness vary across seven commodity groups: food at home (bread, potatoes, vegetables, fruit, meat, dairy products), alcohol, tobacco, food outside the home, clothing, fuel, and luxuries. Moreover, food at home and fuel can be distinguished from alcohol, tobacco, food outside the home, clothing, and luxuries on the basis of a visibility ranking created by

[11] A similar issue arises when the aggregate Q_v^n consists of a low-quality good q_v^l and a high-quality good q_v^h. When a consumer spends significantly more money on the high-quality good, Q_v^n increases with p_v^h. This is a quality effect rather than a diamond effect.

[12] At this point, it is worth noting that we also used an alternative aggregation method. Specifically, we also allowed the weights w_v^k to vary across observations v (following Cherchye et al. (2009)). We found that both methods give similar results. However, fixing the weights w^k is necessary for Hicks' Composite Commodity Theorem to be relevant in our setting.

[13] Lewbel (1996) used a parametric framework to show that correct aggregation is possible even under weaker conditions. Specifically, the generalised aggregation theorem of Lewbel (1996) only requires that the evolution of relative prices (i.e., $p_v^h/p_v^l = \alpha_v$) in an aggregate n is independent of the aggregate price level P_v^n.

Heffetz (2011). To create this ranking, Heffetz (2011) used information from 480 interviews on the visibility of various commodities. The main question in their survey was whether respondents would notice if another household spent more than average on some commodity (e.g., jewellery and watches). Respondents were also asked how much time it would take to notice this greater-than-average spending pattern. In this way, the commodities (not brands) that are most visible to society were expected to obtain a high rank.

Tobacco (ranked 1st), clothing (ranked 3rd), jewellery (ranked 5th), food outside the home (ranked 7th), and alcohol (ranked 8th) are all among the top 10 most visible commodities according to the Heffetz (2011) ranking. Therefore, we assign these commodities to the visible category. Food at home (ranked 14th) and gasoline (ranked 21st) were ranked considerably lower. These commodities are assigned to the invisible category.

In the empirical application, we start from a setting in which all goods (both visible and invisible) are assumed to be standard goods, that is, utility is only derived from the quantities consumed. We investigate whether the behaviour of agents is rational according to the standard GARP test. In the next step, we examine whether (and to what extent) the behaviour of agents can be described by a model that allows for strictly positive diamondness weights. The method also allows us to elicit preferences for value associated with different goods, that is, we allow for heterogeneity across different commodities. Although we do not disentangle direct and indirect preferences for value (i.e., conspicuous consumption), information on the visibility of commodities provides us with an interesting interpretation of the results.

4.5 Application

We first test rationality conditional on various specifications of the diamondness of commodities. We compare specifications on the basis of (average) pass rates, power, and predictive success. We also interpret the results using the visibility index of Heffetz (2011). We expect that stronger preferences for value will correspond to more visible commodities.

Second, we compute predictive success at the individual level. This gives insight into interpersonal heterogeneity in preferences for value. We assess different specifications of diamondness on the basis of the distribution of predictive success (across the sample).

In a final step, we investigate the marginal willingness to pay for value associated with each commodity separately. This gives insight into heterogeneity in diamondness across commodities. Specifically, it is possible to bound the marginal willingness to pay for value using our revealed preference approach.

4.5.1 Testing for rationality with fixed diamondness

Empirical performance Table 4.1 presents the pass rates (average fit of the data) and power estimates associated with different specifications of the model. The rows represent different degrees of diamondness associated with food at home and fuel, and the columns represent different degrees of diamondness associated with alcohol, tobacco, food outside the home, clothing, and luxury commodities. A more detailed decomposition per commodity is provided in Section 4.5.2. Recall that when one particular θ^n equals 0, the respective commodity is valued for its intrinsic consumption component only. By contrast, when θ^n equals 1, the commodity is specifically prised for its value.

The top left result in Table 4.1 corresponds to the neoclassical utility maximisation model (testable with GARP). The behaviour of approximately 56% of the consumers can be rationalised by a well-behaved utility function of the form $U(\mathbf{Q})$. The bottom right result corresponds to the model in which all commodities are prised for their value. Not surprisingly, this revealed preference model imposes no testable restrictions, as a result of which all the data sets are rationalised. The other results are more interesting. By varying the relevant parameters, very different pass rates and (bootstrap) power estimates are obtained.

The predictive success results in Table 4.2 summarise our findings in Table 4.1. We inves-

	diamondness visible goods				
	0	0.25	0.5	0.75	1
diamondness less vis goods					
0	0.561	0.573	0.610	0.634	0.634
	[0.451;0.671]	[0.464;0.682]	[0.502;0.718]	[0.528;0.74]	[0.528;0.74]
	(0.481)	(0.474)	(0.466)	(0.460)	(0.460)
0.25	0.549	0.585	0.610	0.659	0.671
	[0.439;0.659]	[0.476;0.694]	[0.502;0.718]	[0.554;0.764]	[0.567;0.775]
	(0.408)	(0.399)	(0.387)	(0.379)	(0.380)
0.5	0.671	0.720	0.707	0.732	0.793
	[0.567;0.775]	[0.622;0.819]	[0.606;0.808]	[0.634;0.83]	[0.704;0.882]
	(0.326)	(0.310)	(0.293)	(0.280)	(0.279)
0.75	0.732	0.780	0.829	0.866	0.890
	[0.634;0.83]	[0.689;0.871]	[0.746;0.912]	[0.791;0.941]	[0.821;0.959]
	(0.230)	(0.210)	(0.186)	(0.164)	(0.161)
1	0.841	0.890	0.890	0.927	1
	[0.760;0.922]	[0.821;0.959]	[0.821;0.959]	[0.870;0.984]	[1;1]
	(0.145)	(0.113)	(0.076)	(0.039)	(0)

Table 4.1: Mean pass rates, [95 per cent confidence bounds] and (power estimates). less visible goods=fuel and food at home, visible goods=food away from home, clothes, luxuries, tobacco and alcohol

tigate which specifications are empirically supported. The predictive success of the standard model amounts to 0.042. Increasing the diamondness associated with visible consumption generally improves the predictive success results, whereas increasing the diamondness associated with less visible consumption lowers the predictive success scores. In particular, the highest predictive success is obtained when visible commodities have almost full diamondness (i.e., for θ ranging from 0.75 to 1) and the less visible commodities have no diamondness. The corresponding predictive success more than doubles the GARP result.

We finally investigate whether the differences in predictive success are statistically significant. To this end, we apply a procedure described by Demuynck (2014) that allows us to construct 95% confidence bounds around mean predictive success scores. Demuynck (2014) showed how confidence bounds can be computed on the basis of mean pass rates, mean power, and individual pass and power results across the sample. On one hand, the predictive success of neither the standard model nor the specifications giving higher dia-

	diamondness visible goods				
	0	0.25	0.5	0.75	1
diamondness less visible goods					
0	0.042	0.047	0.076	0.094	0.094
	[−0.022; 0.106]	[−0.015; 0.110]	[0.015;0.137]	[0.033;0.155]	[0.033;0.154]
0.25	−0.043	−0.016	−0.003	0.038	0.051
	[−0.102; 0.016]	[−0.073; 0.042]	[−0.059; 0.054]	[−0.017; 0.093]	[−0.004; 0.106]
0.5	−0.003	0.030	0.001	0.012	0.072
	[−0.067; 0.061]	[−0.031; 0.090]	[−0.059; 0.061]	[−0.046; 0.070]	[0.018;0.126]
0.75	−0.039	−0.010	0.015	0.030	0.051
	[−0.101; 0.024]	[−0.070; 0.050]	[−0.040; 0.070]	[−0.019; 0.080]	[0.009;0.094]
1	−0.014	0.003	−0.034	−0.034	0
	[−0.065; 0.037]	[−0.040; 0.046]	[−0.076; 0.008]	[−0.068; −0.001]	[0;0]

Table 4.2: Predictive success (less visible goods=fuel and food at home, visible goods=food away from home, clothes, luxuries, tobacco and alcohol)

mondness to less visible goods is statistically different from 0. On the other hand, the mean predictive success associated with specifications that attribute higher diamondness to visible goods is statistically different from 0. Intuitively, this means that the modified model *can* describe the observed decisions while it rejects most of the random behaviour.

In Appendix 4.C we report power estimates and predictive success results when random bundles are simulated using the Bronars (1987) approach. These results mainly confirm our earlier findings. Once more, the diamondness weights increase with the Heffetz (2011) visibility score. This indicates that consumer preferences for value stem from conspicuous consumption incentives.

Distribution of predictive success Until now, we have investigated the mean predictive success for the sample. We compared different specifications on the basis of average predictive success. It is also insightful to take the distribution of predictive success for the sample into account. After all, we can reasonably argue that consumers are heterogeneous in terms of their preferences for value.

In Figure 4.3 we set out the distribution of predictive success when all goods are standard goods (GARP) and when the visible goods are treated as diamond goods (see supra).

First of all, the mass of individuals whose choices can be described successfully (i.e. with

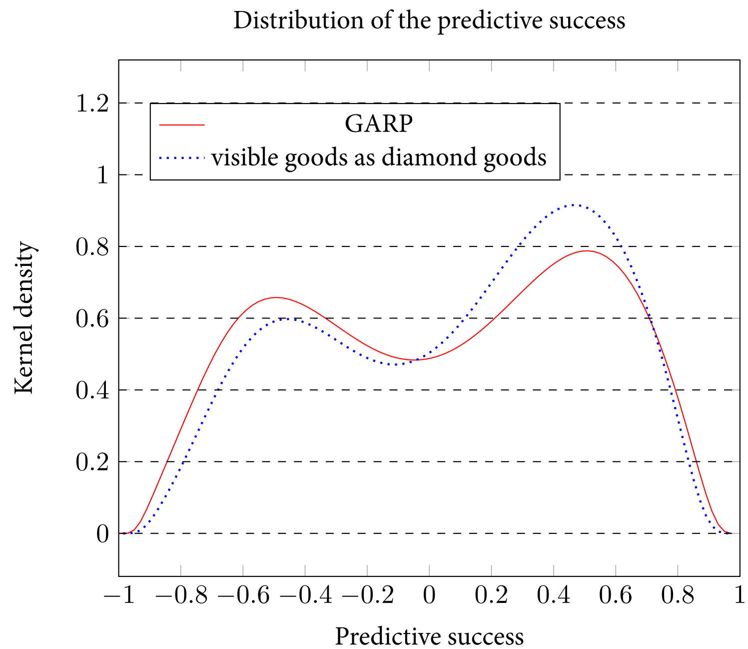


Figure 4.3: Predictive success distribution: standard GARP versus model with visible goods as diamond goods.

a predictive success $>$ 0) is larger than the mass of individuals who are characterised by negative predictive success, for both models. Second, and more interestingly, the predictive success distribution based on the diamondness model has more mass at positive predictive success values. In numbers, the 40th percentile of the predictive success distribution under diamondness is still positive (0.119, meaning that at least 60 per cent of the sample can be described quite reasonably) while the 40th percentile of the distribution under GARP is clearly negative (-0.217). For the sample under consideration, the introduction of diamondness to the visible goods improves the (distribution of) predictive success. The only exception is at the very top of the distribution, where GARP seems to provide slightly more precise results for a very small fraction of the sample.

4.5.2 Identification of diamondness

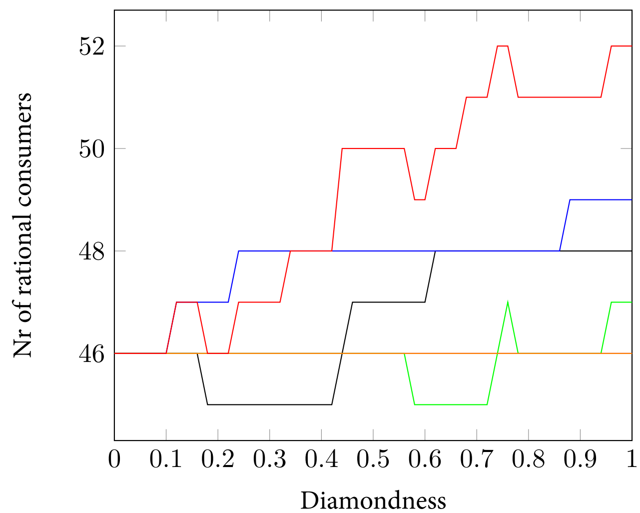
We now go one step further by splitting the visible subgroup into alcohol, tobacco, food away from home, clothing, and luxuries. We analyse if, and to what extent, the number of consumers who behave rationally evolves as a function of diamondness for each product. This also allows us to investigate the heterogeneity in preferences for value across the sample. Indeed, we will be able to 'track' who can (no longer) be rationalised under different specifications of the diamondness vector.

We focus on the visible subgroup for two reasons. First, in the previous subsections we could reject the hypothesis that consumers have strong preferences for value associated with commodities in the *less visible* subgroup. Second, the power estimates in Table 4.1 confirm that discriminatory power is more or less constant for various assumptions on the diamondness of *visible* goods. When analysing visible goods, we can therefore restrict our attention to pass rates.

We set the diamondness of less visible goods to $\theta = 0$ (which is empirically supported by the results in Table 4.2) or $\theta = 0.25$ (for robustness). Table 4.3 describes the change in the number of rational consumers when the diamondness per commodity is increased from 0 to 1 (in steps of 0.1). Figure 4.4 sets out the relationship between the number of rational consumers and the diamondness per commodity graphically. The diamondness of each good is varied unilaterally (denoted by 'alcohol', 'tobacco', 'restaurant', 'clothing' and 'luxuries'). For completeness, we also consider the number of rational consumers as a function of the diamondness of *all* visible goods (denoted by 'all visible').

For the first graph, the diamondness of less visible goods is set to 0. We find that the number of rational consumers improves by 3, 2 and 1, respectively, when the diamondness of alcohol, restaurants and clothing is set to 1. For the second graph, the diamondness of less visible goods is set to 0.25. We find that the number of rational consumers improves by 4, 3, 3 and 3, respectively, when the diamondness of clothing, alcohol, restaurant visits

Pass rates as function of diamondness, $\theta = 0$ for less visible consumption



Pass rates as function of diamondness, $\theta = 0.25$ for less visible consumption

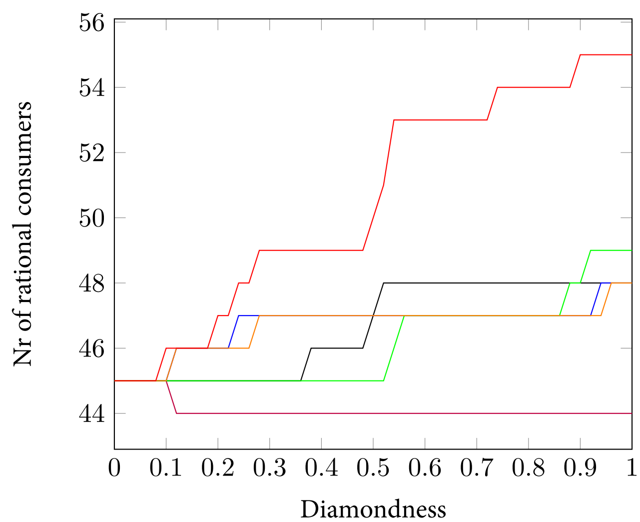


Figure 4.4: Pass rates in function of diamondness

and luxuries is set to 1. This indicates that for our sample alcohol consumption, restaurant expenditures and clothing are the more outspoken diamond goods. We do not find evidence that the consumers in our sample have preferences for value associated with tobacco.

Finally, we consider all visible goods jointly. From Table 4.1 it is clear that 46 consumers are rational (56.1%) when $\theta = 0$, whereas 52 consumers are rational (63.4%) when $\theta = 1$, conditional on all others goods being standard goods. The results in Table 4.3 and Figure 4.4 allow us to investigate this evolution more thoroughly. Specifically, we find that pass rates are not monotonically increasing in the diamondness of one commodity (or one group of commodities). This corresponds to our earlier non-nestedness result: it is possible that decisions are rationalisable given lower diamondness values and not rationalisable for higher diamondness values. For example, at $\theta = 0.169$, the number of individuals that pass the test decreases. Indeed, we can identify a respondent whose marginal willingness to pay for value lies in $[0, 0.169] \cup [0.603, 1]$. Similar observations at $\theta = 0.571$ and $\theta = 0.768$ correspond to respondents whose marginal willingness to pay for value is bounded from above by 0.571 and 0.768, respectively.

θ more vis goods	θ less visible goods = 0						θ less visible goods = 0.25					
	(1)	(2)	(3)	(4)	(5)	(6)	(1)	(2)	(3)	(4)	(5)	(6)
0	0	0	0	0	0	0	0	0	0	0	0	0
0.1	0	0	0	0	0	0	0	0	0	0	0	+1
0.2	-1	0	+1	0	0	0	0	-1	+1	0	+1	+2
0.3	-1	0	+2	0	0	+1	-1	-2	+2	0	+2	+4
0.4	-1	0	+2	0	0	+2	+1	-1	+2	0	+2	+4
0.5	+1	0	+2	0	0	+4	+2	-1	+2	0	+2	+5
0.6	+1	0	+2	-1	0	+3	+3	-1	+2	+2	+2	+8
0.7	+2	0	+2	+1	0	+5	+3	-1	+2	+2	+2	+8
0.8	+2	0	+2	0	0	+5	+3	-1	+2	+2	+2	+9
0.9	+2	0	+3	0	0	+5	+3	-1	+2	+4	+2	+10
1	+2	0	+3	+1	0	+6	+3	-1	+3	+4	+3	+10

Table 4.3: Increase in the number of rational consumers in function of diamondness of visible commodities (columns 1-5: alcohol, tobacco, restaurant, clothing and luxuries, column 6: all visible goods jointly)

4.6 Conclusion

We incorporated preferences for value in the revealed preference framework. We defined *diamondness* as the marginal willingness to pay for the value associated with commodities. Strictly positive diamondness implies that consumers derive utility from the value of goods, not just from the quantity consumed.

We first generalised the model of Ng (1993) by allowing for more than one diamond good. Moreover, we let the degree of diamondness vary on a scale from 0 to 1. Interestingly, the newly proposed diamondness parameter measures the diamond effect in monetary terms. By extending the neoclassical model of rationality with one (set of) parameter(s), our approach fits the PEEM research agenda supported by Rabin (2013).

Next, we constructed revealed preference conditions that can be used to verify assumptions on the diamondness of commodities. An attractive feature of the revealed preference methodology is that it imposes minimal restrictions on the form of utility functions. Moreover, each agent can be analysed separately so that debatable preference homogeneity assumptions can be avoided. We showed that the revealed preference approach produces refutable conditions even if preferences depend on value. In this respect, we also argued that the different characterisations (corresponding to different assumptions for the diamondness vector) are generally non-nested. The method can produce meaningful bounds on the diamondness associated with particular commodities for particular consumers.

Finally, we applied our nonparametric test of preferences for value to a data sample from the RLMS. To the best of our knowledge, this is the first application of revealed preference that specifically incorporates preferences for value. We found that the predictive success of a standard GARP test is significantly less than that of alternative specifications that set strictly positive marginal willingness to pay for value. To interpret our results, we investigated the relationship between marginal willingness to pay for value and the visibility of a commodity to society. On the one hand, the hypothesis that less visible commodities have strong diamondness was rejected. It is unlikely that individuals have preferences for value associated with food consumed at home and fuel. This might explain why Heffetz and Shayo (2009), who focused on food, did not find significant non-budget-constraint price effects in their experiment. On the other hand, our results suggest that consumers have strong preferences for value associated with more visible commodities, such as clothing. Following the argument of Ng (1987), special taxation rules may be appropriate. Furthermore, we found much variation in the rationality results across the sample. This clearly shows that choices and preferences are very diverse, and that homogeneity assumptions on preferences for value are probably unrealistic. Finally, we set out rationality results for all respondents as a function of the diamondness of alcohol, tobacco, food away from home, clothing, and luxuries. We found that pass rates generally increase with the diamondness of all visible commodities apart from tobacco. For particular individuals, we could also establish meaningful bounds on their marginal willingness to pay for value.

The main contribution of the current study is that it captures the extent to which value-dependent preferences (diamondness defined in a broad sense) are important. The question remains whether the diamond effect (in line with Ng (1987)) and the closely related status effect (conspicuous consumption) are distinguishable from exogenous quality changes within aggregates and from exogenous quality changes across aggregates. The answer to the first question depends on whether Hicks' Aggregation Theorem holds. The results are robust to exogenous (within-aggregate) changes in quality or composition as long as the relative prices within the aggregates are sufficiently stable over time. The answer to the second question depends on the type of application under consideration. In experiments, it is possible to set up choice problems in which quality and status effects are distinguishable. Such setting could allow researchers to elicit the precise motivations underlying the preferences for value. In observational budget surveys, it seems difficult to discriminate between pure diamond effects and quality changes. In this respect, however, it is worth noting that consumption changes as a result of aggregate quality changes (e.g. "I buy more food because the quality of food has gone up") are less likely. Finally, to examine the robustness and external validity of our findings, it would be interesting to see applications of the model to larger samples, and to collective settings in which various other factors (such as the affection between two partners) impact on the consumption of diamond goods.

4.A Proof of Proposition 4.4

Proof. In this appendix, we present the proof of Proposition 4.4.

- We first prove that Condition 1 implies Condition 2. Consider the following (necessary) first-order condition for the optimisation of $OPT-\theta$:

$$\partial U/\partial Q_t^n + \partial U/\partial M_t^n \cdot P_t^n \leq \lambda_t P_t^n;$$

  where $\partial U/\partial Q_t^n$ and $\partial U/\partial (P_t^n Q_t^n)$ are subderivatives of the (concave) utility function with respect to Q_t^n and $P_t^n Q_t^n$, respectively. The inequalities are replaced with equalities if the quantities Q_t^n are strictly positive.

  Moreover, concavity of the utility function gives:

$$u_t - u_v \leq \sum_{n=1}^{|N|} \partial U/\partial Q_v^n \cdot (Q_t^n - Q_v^n) + \sum_{n=1}^{|N|} \partial U/\partial M_v^n \cdot (P_t^n Q_t^n - P_v^n Q_v^n); \quad (4.6)$$

  Finally, we replace $\mathfrak{p}_v^n = \dfrac{\partial U/\partial Q_v^n}{\lambda_v}$ and $\mathfrak{P}_v^n = 1 - \dfrac{\mathfrak{p}_v^n}{P_v^n}$ such that $\mathfrak{P}_v^n \geq \dfrac{\partial U/\partial M_v^n}{\lambda_v}$. For strictly positive quantities Q^n, we have that $\mathfrak{P}_v^n = \dfrac{\partial U/\partial M_v^n}{\lambda_v}$. Consistency with Condition (4.6) thus requires consistency with Condition (4.7).

$$u_t - u_v \leq \lambda_v \sum_{n=1}^{|N|} \mathfrak{p}_v^n \cdot (Q_t^n - Q_v^n) + \lambda_v \sum_{n=1}^{|N|} \mathfrak{P}_v^n \cdot (P_t^n Q_t^n - P_v^n Q_v^n); \qquad (4.7)$$

$$P_v^n = \mathfrak{p}_v^n + \mathfrak{P}_v^n \cdot P_v^n \quad \forall n \in N;$$

This concludes the *necessity* part.

- We then prove that Condition 2 implies Condition 1 (based on Varian (1982)).

  We start from the observation that

$$U(\mathbf{Q},\mathbf{M}) \leq U_v + \lambda_v \sum_{n=1}^{|N|} \mathfrak{p}_v^n \cdot (Q^n - Q_v^n) + \lambda_v \sum_{n=1}^{|N|} \mathfrak{P}_v^n \cdot (P^n Q^n - P_v^n Q_v^n);$$

  In a following step, we select the minimum of all overestimates:

$$U(\mathbf{Q},\mathbf{M}) = \min_v (U_v + \lambda_v \sum_{n=1}^{|N|} \mathfrak{p}_v^n \cdot (Q^n - Q_v^n) + \lambda_v \sum_{n=1}^{|N|} \mathfrak{P}_v^n \cdot (P^n Q^n - P_v^n Q_v^n));$$

  This formulation should be such that any $(\mathbf{Q},\mathbf{M})$ for which $\mathbf{P}_t'\mathbf{Q}_t \geq \mathbf{P}_t'\mathbf{Q}$, implies that $U(\mathbf{Q}_t,\mathbf{M}_t) \geq U(\mathbf{Q},\mathbf{M})$.

  First, it is important to understand that $U(\mathbf{Q}_v,\mathbf{M}_v) = U_v$ for $v = 1, ..., T$. Indeed, for some t, we have that

$$\begin{aligned} U(\mathbf{Q}_v,\mathbf{M}_v) &= U_t + \lambda_t \sum_{n=1}^{|N|} \mathfrak{p}_t^n \cdot (Q_v^n - Q_t^n) + \lambda_t \sum_{n=1}^{|N|} \mathfrak{P}_t^n \cdot (P_v^n Q_v^n - P_t^n Q_t^n) \\ &\leq U_v + \lambda_v \sum_{n=1}^{|N|} \mathfrak{p}_v^n \cdot (Q_v^n - Q_v^n) + \lambda_v \sum_{n=1}^{|N|} \mathfrak{P}_v^n \cdot (P_v^n Q_v^n - P_v^n Q_v^n) \\ &= U_v \end{aligned}$$

  If this inequality were strict, we would have that

$$U_v - U_t > \lambda_t \sum_{n=1}^{|N|} \mathfrak{p}_t^n \cdot (Q_v^n - Q_t^n) + \lambda_t \sum_{n=1}^{|N|} \mathfrak{P}_t^n \cdot (P_v^n Q_v^n - P_t^n Q_t^n)$$

  which contradicts the Afriat inequalities. Hence $U(\mathbf{Q}_v,\mathbf{M}_v) = U_v$.

Second, any $(\mathbf{Q}, \mathbf{M})$ for which $\mathbf{P}_t'\mathbf{Q}_t \geq \mathbf{P}_t'\mathbf{Q}$ must be consistent with

$$
\begin{aligned}
U(\mathbf{Q}, \mathbf{M}) &= \min_v (U_v + \lambda_v \sum_{n=1}^{|N|} \mathfrak{p}_v^n \cdot (Q^n - Q_v^n) + \lambda_v \sum_{n=1}^{|N|} \mathfrak{P}_v^n \cdot (P_t^n Q^n - P_v^n Q_v^n)) \\
&\leq U_t + \lambda_t \sum_{n=1}^{|N|} \mathfrak{p}_t^n \cdot (Q^n - Q_t^n) + \lambda_t \sum_{n=1}^{|N|} \mathfrak{P}_t^n \cdot (P_t^n Q^n - P_t^n Q_t^n) \\
&\leq U_t = U(\mathbf{Q}_t, \mathbf{M}_t)
\end{aligned}
$$

The first inequality follows from the definition of $U(\mathbf{Q}, \mathbf{M})$, the second inequality follows from

$$
\begin{aligned}
& U_t + \lambda_t \sum_{n=1}^{|N|} \mathfrak{p}_t^n \cdot (Q^n - Q_t^n) + \lambda_t \sum_{n=1}^{|N|} \mathfrak{P}_t^n \cdot (P_t^n Q^n - P_t^n Q_t^n) \\
= \; & U_t + \lambda_t \sum_{n=1}^{|N|} (P_t^n - \mathfrak{P}_t^n \cdot P_t^n) \cdot (Q^n - Q_t^n) + \lambda_t \sum_{n=1}^{|N|} \mathfrak{P}_t^n \cdot (P_t^n Q^n - P_t^n Q_t^n) \\
= \; & U_t + \lambda_t \sum_{n=1}^{|N|} P_t^n \cdot (Q^n - Q_t^n) \\
\leq \; & U_t
\end{aligned}
$$

This concludes the *sufficiency* part.

$\square$

4.B Hicks' Composite Commodity Theorem in the diamond setting

In this proof, we show that Hicks' Composite Commodity Theorem also applies in a setting with diamond goods. Otherwise stated, the testable conditions in Proposition 4.4 give the same results, regardless of whether the restrictions are applied to the aggregates with unit values or to the subgroups with specific prices, provided that there is no relative price variation within the aggregates.

Proof. We start from aggregate prices which are constructed in the following way

$$P_v^n = (p_v^l)^w (p_v^h)^{1-w}$$

and from the assumption that there is no (relative) price variation within subgroups of products, i.e. $\frac{p_v^h}{p_v^l} = \alpha$.

- First, we show that this information enables us to express Q_v, p_v^l and p_v^h in terms of constants $\beta^l = \alpha^{w-1}$ and $\beta^h = \alpha^w$, which are invariant across periods v.

$$\begin{aligned} Q_v &= \beta^l q_v^l + \beta^h q_v^h \\ p_v^l &= \beta^l P_v \\ p_v^h &= \beta^h P_v \end{aligned}$$

In order to see this, we use that $\frac{p_v^h}{p_v^l} = \alpha$ and $P_v^n = (p_v^l)^w (p_v^h)^{1-w}$:

$$\begin{aligned} Q_v &= \frac{p_v^l q_v^l + p_v^h q_v^h}{P_v^n} = \frac{p_v^l q_v^l + p_v^h q_v^h}{(p_v^l)^w (p_v^h)^{1-w}} = \frac{p_v^l q_v^l + \alpha p_v^l q_v^h}{(p_v^l)^w (\alpha p_v^l)^{1-w}} \\ &= \frac{p_v^l (q_v^l + \alpha q_v^h)}{p_v^l (\alpha^{1-w})} = \frac{q_v^l + \alpha q_v^h}{\alpha^{1-w}} = \alpha^{w-1} q_v^l + \alpha^w q_v^h = \beta^l q_v^l + \beta^h q_v^h \end{aligned}$$

$$\begin{aligned} P_v &= (p_v^l)^w (p_v^h)^{1-w} = (p_v^l)^w (\alpha p_v^l)^{1-w} = \alpha^{1-w} p_v^l \\ &\Rightarrow \; p_v^l = \beta^l P_v \end{aligned}$$

$$\begin{aligned} P_v &= (p_v^l)^w (p_v^h)^{1-w} = (\frac{p_v^h}{\alpha})^w (p_v^h)^{1-w} = \alpha^{-w} p_v^h \\ &\Rightarrow \quad p_v^h = \beta^h P_v \end{aligned}$$

- Next, reformulate the first-order conditions for all subgroups $k \in \{h, l\}$ that belong to the aggregate n :

$$\begin{aligned} p_v^{kn} &= \mathfrak{p}_v^{kn} + \mathfrak{P}_v^{kn} \cdot p_v^{kn} \\ &\Leftrightarrow \quad \beta^{kn} P_v^n = \mathfrak{p}_v^{kn} + \mathfrak{P}_v^{kn} \cdot \beta^{kn} P_v^n \\ &\Leftrightarrow \quad P_v^n = \frac{\mathfrak{p}_v^{kn}}{\beta^{kn}} + \mathfrak{P}_v^{kn} \cdot P_v^n \end{aligned}$$

Hence, we can conclude that, for fixed $\theta^n = \theta^{ln} = \theta^{hn}$ such that $\mathfrak{P}_v^n = \mathfrak{P}_v^{ln} = \mathfrak{P}_v^{hn}$:

$$\frac{\mathfrak{p}_v^{hn}}{\beta^{hn}} = \frac{\mathfrak{p}_v^{ln}}{\beta^{ln}}$$

We can simply redefine $\mathfrak{p}_v^n = \frac{\mathfrak{p}_v^{hn}}{\beta^{hn}} = \frac{\mathfrak{p}_v^{ln}}{\beta^{ln}}$.

- Finally, we can show the equivalence between the inequalities in Proposition 4.4 (Statement 2) applied to the subproducts on the one hand and the inequalities in Proposition 4.4 (Statement 2) applied to the aggregates on the other hand.

$$
\begin{aligned}
u_t - u_v \quad \leq \quad & \lambda_v \mathfrak{p}_v^{h1}(q_t^{h1} - q_v^{h1}) + \lambda_v \mathfrak{p}_v^{h2}(q_t^{h2} - q_v^{h2}) \\
& + \lambda_v \mathfrak{p}_v^{l1}(q_t^{l1} - q_v^{l1}) + \lambda_v \mathfrak{p}_v^{l2}(q_t^{l2} - q_v^{l2}) \\
& + \lambda_v \mathfrak{P}_v^{1}(p_t^{h1} q_t^{h1} - p_v^{h1} q_v^{h1}) + \lambda_v \mathfrak{P}_v^{2}(p_t^{h2} q_t^{h2} - p_v^{h2} q_v^{h2}) \\
& + \lambda_v \mathfrak{P}_v^{1}(p_t^{l1} q_t^{l1} - p_v^{l1} q_v^{l1}) + \lambda_v \mathfrak{P}_v^{2}(p_t^{l2} q_t^{l2} - p_v^{l2} q_v^{l2})
\end{aligned}
$$

$$
\begin{aligned}
\Leftrightarrow \quad & u_t - u_v \leq \lambda_v \mathfrak{p}_v^{1}(\beta^{h1} q_t^{h1} - \beta^{h1} q_v^{h1}) + \lambda_v \mathfrak{p}_v^{2}(\beta^{h2} q_t^{h2} - \beta^{h2} q_v^{h2}) \\
& + \lambda_v \mathfrak{p}_v^{1}(\beta^{l1} q_t^{l1} - \beta^{l1} q_v^{l1}) + \lambda_v \mathfrak{p}_v^{2}(\beta^{l2} q_t^{l2} - \beta^{l2} q_v^{l2}) \\
& + \lambda_v \mathfrak{P}_v^{1}(P_t^1 \beta^{h1} q_t^{h1} - P_v^1 \beta^{h1} q_v^{h1}) + \lambda_v \mathfrak{P}_v^{2}(P_t^2 \beta^{h2} q_t^{h2} - P_v^2 \beta^{h2} q_v^{h2}) \\
& + \lambda_v \mathfrak{P}_v^{1}(P_t^1 \beta^{l1} q_t^{l1} - P_v^1 \beta^{l1} q_v^{l1}) + \lambda_v \mathfrak{P}_v^{2}(P_t^2 \beta^{l2} q_t^{l2} - P_v^2 \beta^{l2} q_v^{l2})
\end{aligned}
$$

$$
\begin{aligned}
\Leftrightarrow \quad & u_t - u_v \leq \lambda_v \mathfrak{p}_v^{1}(\beta^{h1} q_t^{h1} + \beta^{l1} q_t^{l1} - \beta^{h1} q_v^{h1} - \beta^{l1} q_v^{l1}) \\
& + \lambda_v \mathfrak{p}_v^{2}(\beta^{h2} q_t^{h2} + \beta^{l2} q_t^{l2} - \beta^{h2} q_v^{h2} - \beta^{l2} q_v^{l2}) \\
& + \lambda_v \mathfrak{P}_v^{1}(P_t^1 \beta^{h1} q_t^{h1} + P_t^1 \beta^{l1} q_t^{l1} - P_v^1 \beta^{h1} q_v^{h1} - P_v^1 \beta^{l1} q_v^{l1}) \\
& + \lambda_v \mathfrak{P}_v^{2}(P_t^2 \beta^{h2} q_t^{h2} + P_t^2 \beta^{l2} q_t^{l2} - P_v^2 \beta^{h2} q_v^{h2} - P_v^2 \beta^{l2} q_v^{l2})
\end{aligned}
$$

$$
\begin{aligned}
\Leftrightarrow \quad & u_t - u_v \leq \lambda_v \mathfrak{p}_v^{1}(Q_t^1 - Q_v^1) + \lambda_v \mathfrak{p}_v^{2}(Q_t^2 - Q_v^2) \\
& + \lambda_v \mathfrak{P}_v^{1}(P_t^1 Q_t^1 - P_t^1 Q_v^1) + \lambda_v \mathfrak{P}_v^{2}(P_t^2 Q_t^2 - P_v^2 Q_v^2)
\end{aligned}
$$

$\square$

4.C Predictive success based on Bronars' method

For robustness, we report power and predictive success measures based on Bronars' method.

	diamondness visible goods				
	0	0.25	0.5	0.75	1
diamondness less visible goods					
0	0.561	0.573	0.610	0.634	0.634
	(0.452)	(0.445)	(0.442)	(0.440)	(0.443)
0.25	0.549	0.585	0.610	0.659	0.671
	(0.379)	(0.366)	(0.358)	(0.354)	(0.357)
0.5	0.671	0.720	0.707	0.732	0.793
	(0.303)	(0.287)	(0.272)	(0.264)	(0.262)
0.75	0.732	0.780	0.829	0.866	0.890
	(0.219)	(0.191)	(0.166)	(0.150)	(0.147)
1	0.841	0.890	0.890	0.927	1
	(0.160)	(0.126)	(0.080)	(0.037)	(0)

Table 4.4: Pass rates (and Bronars' power estimates) (less visible goods=fuel and food at home, visible goods=food away from home, clothes, luxuries, tobacco and alcohol)

Most of our earlier conclusions remain valid. First of all, the predictive success scores generally increase in function of the diamondness for visible goods.

None of the predictive success scores are statistically different from 0 at the 5 per cent level. At the 10 per cent level, however, we find that (only) some characterisations, which attribute high diamondness to the visible goods, have a predictive success score that is statistically different from 0.

	diamondness visible goods				
	0	0.25	0.5	0.75	1
diamondness less visible goods					
0	0.013	0.018	0.052	0.075	0.077
	[-0.050;0.076]	[-0.043;0.078]	[-0.008;0.112]	[0.015;0.134]	[0.016;0.137]
0.25	-0.073	-0.048	-0.033	0.012	0.027
	[-0.131;-0.014]	[-0.105;0.009]	[-0.089;0.024]	[-0.044;0.068]	[-0.028;0.083]
0.5	-0.026	0.007	-0.020	-0.005	0.055
	[-0.089;0.036]	[-0.052;0.065]	[-0.079;0.038]	[-0.062;0.053]	[0.002;0.108]
0.75	-0.050	-0.029	-0.005	0.016	0.037
	[-0.112;0.013]	[-0.087;0.030]	[-0.059;0.049]	[-0.034;0.065]	[-0.007;0.080]
1	0.001	0.016	-0.029	-0.036	0
	[-0.049;0.051]	[-0.025;0.057]	[-0.071;0.012]	[-0.070;-0.001]	[0;0]

Table 4.5: Predictive success based on Bronars' method (less visible goods=fuel and food at home, visible goods=food away from home, clothes, luxuries, tobacco and alcohol)

Part IV

Bounds on the distribution of welfare and demand in a heterogeneous population

In this part, I combine elementary revealed preference principles and nonparametric estimation techniques to obtain nonparametric bounds on the distribution of the money metric utility over a population of heterogeneous households. The method can also produce bounds on demand predictions in counterfactual price-income regimes.

I build on the contributions of Blundell et al. (2003, 2007, 2008). These authors have shown how to use the revealed preference axioms, in combination with kernel estimates of Engel functions, to bound welfare and demand for a representative consumer. Their contributions addressed Problems 1 and 2 from my General Introduction. First of all, the method can be applied to cross-sectional data sets. This considerably enlarges the scope of revealed preference. Second, it produces reasonably tight bounds on welfare estimates and demand correspondences.

However, the approach presented in this part focuses on the entire distribution of welfare and demand, rather than the welfare and demand corresponding to a representative agent. There is plenty of evidence that preferences and tastes are heterogeneous across consumers. This requires us to incorporate very general forms of unobserved heterogeneity. After all, Lewbel (2001) has pointed out that imposing additive error terms comes very close to the representative agent assumption. To include very general forms of unobserved heterogeneity, I build on the stochastic revealed preference approach developed by McFadden and Richter (1971) and Falmagne (1978).

Chapter 5

Nonparametric bounds for a heterogeneous population[1]

5.1 Introduction

We present a framework to construct nonparametric bounds on the distribution of the money metric utility function while taking into account individual unobserved heterogeneity. Our approach combines elementary revealed preference concepts (in particular the Weak Axiom of Revealed Preference) with nonparametric (kernel) estimation techniques. In this manner, our approach remains independent of any parametric specification on the underlying household utility functions or on the unobserved heterogeneity distribution. We further demonstrate how the framework can be used to establish bounds on the distribution of the demand functions in counterfactual price regimes. An illustration using the Consumer Expenditure Survey, a US cross–sectional household consumption data set, demonstrates the practical usefulness of our results.

[1]This chapter is based on joint work with Thomas Demuynck (Maastricht University). I refer to the working paper version of Cosaert and Demuynck (2014a).

Motivation Demand analysis provides a powerful tool to analyse behavioural responses and welfare effects due to price and income variations. In a typical demand study, the researcher first estimates the parameters of some parametric demand system,[2] and uses these estimates to calculate the associated indirect utilities. This 'parametric' approach has two major shortcomings. The first shortcoming is that the outcome is sensitive to the specific functional structure chosen by the researcher. Imposing the wrong functional form can therefore severely bias the resulting analysis. A second shortcoming concerns the treatment of individual (unobserved) heterogeneity. In a typical consumer data set, we observe individuals or households only once. Given this data limitation, it is often assumed that similar looking individuals have similar preferences. Many demand studies therefore model a household's demand to equal a rational systematic component, from a common utility function across all (similar looking) households, and a household specific additive error term capturing the unobserved heterogeneity or taste variation. By controlling for various observable characteristics (like household size), it is hoped that the issue of heterogeneity across the households is adequately addressed by including such additive error term. This assumption, however, disregards the finding that individuals who look very similar may actually differ dramatically in their actual choice behaviour.[3] As shown by Lewbel (2001), imposing additivity of the unobserved heterogeneity is a strong assumption. Its resulting implications come very close to enforcing a representative agent assumption.[4] To summarise, we see that different people (although they may look the same) have different tastes and, consequentially, behave differently. In order to take this into account, it is crucial to allow for non-additive unobserved heterogeneity.

[2] Popular parametric demand systems are the Translog (Christensen, Jorgenson, and Lau (1975)), the Almost Ideal (Deaton and Muellbauer (1980)), or the Quadratic Almost Ideal (Banks, Blundell, and Lewbel (1997)) demand system.

[3] Unobserved heterogeneity is often seen as the main reason why demand estimations on cross sectional data typically have low r-squared values.

[4] See also Brown and Walker (1989) and McElroy (1987) for a discussion of other issues when taking into account unobserved heterogeneity.

Literature overview In order to deal with aforementioned two problems, one can distinguish between two approaches. The first approach looks at the nonparametric differential 'smooth' restriction that can still be established in a heterogeneous population. These usually take the form of population level generalisations of Slutsky symmetry, negativity and homogeneity. Recent examples that follow this approach are Hoderlein (2011), Blundell, Horowitz, and Parey (2013), Hausman and Newey (2013), and Hoderlein and Vanhems (2013). A second approach, followed in this paper, is to rely on revealed preference theory. Revealed preference theory was initiated by Samuelson (1938), Houthakker (1950) and further developed in several seminal contributions by Afriat (1967), Diewert (1973) and Varian (1982). The main aim of revealed preferences theory is to establish (combinatorial) restrictions on observed demand behaviour of a certain individual or household such that it is consistent with the classical model of utility maximisation subject to a budget constraint. One of the main advantages of revealed preference theory is that it imposes no functional restrictions on the underlying utility function, except for some regularity conditions like local non-satiation.

Revealed preference theory, as it was initially developed, has two main problems. First, from an empirical point of view, the method does not really seem to provide very tight bounds. The main reason for this is that relative price variations usually tend to be quite small in comparison to income variation. This implies that budget hyperplanes often do not cross. We refer to Bronars (1987) and Varian (1982) for a discussion of this problem. The second problem is that revealed preference theory is not well suited to deal with unobserved individual heterogeneity. As a result, most of its applications remain confined to a few panel consumption data sets, where the same household or individual is observed over multiple periods.

The first problem has been the subject of several recent studies that apply revealed preference theory to repeated cross sectional data by combining insights from revealed preference theory with nonparametric estimation techniques (see Blundell (2005); Blundell, Browning, and Crawford (2003, 2007, 2008) and Blundell, Browning, Cherchye, Crawford, De Rock, and Vermeulen (2012)). The main contribution from this literature is that it shows how to use nonparametric Engel curve demand estimates as an input for revealed preference analysis. If we assume that households in the same time period and location face the same relative prices, then the nonparametric Engel curves estimate the mean (or average) expansion paths for each price regime. The availability of these expansion paths greatly improves the nonparametric bounds on various welfare related concepts and on the counterfactual demand estimates that can be obtained using revealed preference techniques.

A remaining drawback of this approach is the way it deals with the issue of unobserved heterogeneity. Given that the Engel curve estimates are obtained from a mean regression, the methodology is subject to Lewbel (2001)'s critique: imposing revealed preference restrictions on the mean Engel curve estimates comes very close to imposing a representative consumer assumption. Given this, the approach does not fully address the individual heterogeneity problem. Moreover, despite the fact that the procedure has the potential to produce tight bounds on the 'representative' money metric utility and demand functions, it does not give us any information concerning the distribution of these functions across the heterogeneous population.

A useful extension of revealed preference theory that explicitly takes into account individual heterogeneity is Stochastic Revealed Preference Theory, initiated by McFadden and Richter (1971) and Falmagne (1978).[5] We refer to McFadden (2005) for an overview of the literature. Stochastic revealed preference takes as input the entire distribution of demand behaviour over a heterogeneous population of households for a finite number of budget sets.[6] Therefore, it is well suited to deal with the issue of unobserved heterogeneity. The literature has put forward several rationality axioms (e.g. the Axiom of Revealed Stochastic Preference

[5]See also Block and Marschak (1959), McFadden (1975), Fishburn (1978), Cohen (1980), Barberá and Pattanaik (1986), Fishburn and Falmagne (1989), Cohen and Falmagne (1990), Fishburn (1992) and Bandyopadhyay, Dasgupta, and Pattanaik (1999) for other contributions.

[6]A second interpretation of stochastic revealed preference theory is that the demand behaviour is generated by a single household with a random utility function.

and the Weak Axiom of Stochastic Revealed Preference) that provide conditions on the distributions of choices such that a population of individuals is consistent with rational choice theory, which postulates that individuals are preference maximisers. Although the literature is mainly theoretical, several recent papers have started to develop statistical tests to verify whether the stochastic revealed preference axioms are satisfied in reality. Hoderlein and Stoye (2014) derive a statistical procedure to infer bounds on the fraction of the population that violates the Weak Axiom of Stochastic Revealed Preference. Kitamura and Stoye (2013) derive a statistical test to verify whether a population of heterogeneous households satisfies the Axiom of Stochastic Revealed Preference for a finite collection of budget sets, thereby explicitly taking into account that preferences are transitive. Finally, Kawaguchi (2012) derives several procedures to test the validity of various axioms of revealed stochastic preference. Interestingly, these studies find little evidence that the stochastic revealed preference restrictions are violated. The main difference between these papers and ours is the focus. While the existing contributions mainly deal with testing whether the axioms imposed by the stochastic revealed preference literature hold, we are more interested in the restrictions that the stochastic revealed preference axioms impose on the resulting distribution of the money metric utility and demand functions. In the terminology of Varian (1982); while above papers deal with *testing* the theory, we concentrate on the *recovery* of the underlying structure of the model.

Another closely related paper is Blundell, Kristensen, and Matzkin (2014). These authors focus on the issue of unobserved heterogeneity in a two goods setting. In particular, they tackle the problem of individual unobserved heterogeneity using nonparametric quantile demand estimates in combination with standard revealed preference tests (i.e. SARP). Hoderlein and Stoye (2013) recently showed that in a two goods setting, imposing the usual revealed preference axioms on the quantile demands is equivalent to imposing the Axiom of Stochastic Revealed Preference on the entire data set.[7] Moreover, the analysis of Blun-

[7] In a two-goods setting, the analysis is simplified by the fact that in a two goods setting, the Weak Axiom of

dell, Kristensen, and Matzkin (2014) is based on (at least) two assumptions: the random component is univariate and uniformly distributed and demand is invertible in the random component. In our framework, we abstain from imposing these assumptions.

Contribution The main contribution of this paper is to derive nonparametric bounds on the money metric utility functions and the demand functions without imposing any functional structure on the household utility functions and the unobserved heterogeneity structure. As such, we avoid the problem that our results might be biased because of a wrong functional specification or because the households do not satisfy the 'representative agent' condition. We establish our results by combining elementary stochastic revealed preference theory and nonparametric estimation techniques. Our framework not only allows us to derive bounds on the mean of the money metric utility and demand functions, but on the entire distribution of these functions over the heterogeneous population. This provides important additional information concerning the distribution of welfare and demand over the population.

In order to obtain our results, we exploit the Weak Axiom of Revealed Preferences (WARP) applied to a population of heterogeneous households. Although this axiom is weaker than the revealed preference axioms that exploit transitivity (e.g. the Strong Axiom of Revealed Preference), we nevertheless show that it is powerful enough to establish narrow bounds. We demonstrate the usefulness of our results by applying it to the Consumer Expenditure survey, a US cross sectional consumption data set.

In Section 5.2, we set out our framework and we present the necessary notation, concepts and definitions for the remaining part of the paper. Section 5.3 establishes the theoretical results that provide the bounds on the distribution of the money metric utility function and the demand functions. Section 5.4 contains our empirical application. We discuss estimation, statistical inference and we present several results. Section 5.5 concludes and points

(Stochastic) Revealed Preference coincides with the Strong Axiom of (Stochastic) Revealed Preference. In other words, imposing transitivity implies no additional testable implications, see Rose (1958).

towards future research.

5.2 Notation and Definitions

In this section, we set out our basic framework and we introduce the notation and definitions that are needed in order to establish the results in the following sections.

Set up We consider an economy with a large (infinite) number of different households. Each household, h, is endowed with a utility function which we denote by $u(\mathbf{q}^h; \mathbf{a}^h, \boldsymbol{\nu}^h)$. This utility function depends on a (column) vector of consumed goods $\mathbf{q}^h \in \mathbb{R}^n_+$, where n is the number of goods, a vector of observable household specific attributes $\mathbf{a}^h$, e.g. household composition, and a household specific vector of unobservable attributes $\boldsymbol{\nu}^h$, capturing unobserved preference heterogeneity.[8] In order to decide how much to consume, the household maximises its utility function subject to a household budget constraint,

$$\mathbf{q}(\mathbf{p}, x^h; \mathbf{a}^h, \boldsymbol{\nu}^h) = \arg\max_{\mathbf{q}} u(\mathbf{q}; \mathbf{a}^h, \boldsymbol{\nu}^h) \text{ s.t. } \mathbf{pq} \leq x^h.$$

Here we denote by $\mathbf{p} \in \mathbb{R}^n_{++}$ a (row) vector of strictly positive prices and by $x^h \in \mathbb{R}_{++}$ the total household disposable income. We assume that the solution of this optimisation problem gives a system of n demand functions $\mathbf{q}^h = \mathbf{q}(\mathbf{p}, x^h; \mathbf{a}^h, \boldsymbol{\nu}^h)$ which depend on the vector of prices, the income and the household observable and unobservable attributes. We assume that the utility function is strictly quasi-concave and twice continuously differentiable in $\mathbf{q}$ such that the demand functions are well defined and continuous in $\mathbf{p}$ and x. For a price vector $\mathbf{p}_t$ and an expenditure level x, we denote by $(\mathbf{p}_t, x)$, the budget set consisting of all bundles $\mathbf{q}$ such that $\mathbf{p}_t\mathbf{q} \leq x$.

We treat $\boldsymbol{\nu}$ as a random vector. Using $F(\boldsymbol{\nu}|\cdot)$ to denote the conditional distribution

[8] Different households have different utility functions because they have different values of the vectors $\mathbf{a}^h$ and $\boldsymbol{\nu}^h$. In other words, it is as if each household has a household specific utility function $u^h(.) = u(.; \mathbf{a}^h, \boldsymbol{\nu}^h)$.

of the unobserved preference attributes over the population of households, we impose the following assumption.

Assumption 5.1. For all income levels x and prices $\mathbf{p}$,

$$F(\boldsymbol{\nu}|\mathbf{p}, x, \mathbf{a}) = F(\boldsymbol{\nu}|\mathbf{a}).$$

Assumption 5.1 requires that the vector of unobserved attributes is independent of prices and income, conditional on all observable attributes. This 'independence of budget sets' condition is frequently used in the literature.[9] If we interpret $\boldsymbol{\nu}$ as a vector of preference parameters, Assumption 5.1 encompasses the idea, common in consumer demand, that preferences do not vary with prices and income. For notational convenience, we omit from now the dependence on the observable attributes $\mathbf{a}$, taking into account that every expression is valid conditional on a particular value of this vector.

For the remaining part of the paper, it will be more useful to work with the indirect utility function $v(\mathbf{p}, x^h; \boldsymbol{\nu}^h)$ which gives the maximal utility that household h can obtain at prices $\mathbf{p}$ and income x^h. The indirect utility function is defined from the direct utility function by,

$$v(\mathbf{p}, x^h; \boldsymbol{\nu}^h) = u(\mathbf{q}(\mathbf{p}, x^h; \boldsymbol{\nu}^h); \boldsymbol{\nu}^h).$$

The indirect utility function is strictly increasing in the level of disposable income x^h. If we invert the indirect utility function $v(\mathbf{p}, x^h; \boldsymbol{\nu}^h)$, with respect to x^h, we obtain the expenditure function $e(\mathbf{p}, u^h; \boldsymbol{\nu}^h)$ which gives the minimal outlay for household h to reach utility level u^h at prices $\mathbf{p}$. Finally, using the expenditure function, we can define the money metric utility function,

$$\mu(\mathbf{p}_v, \mathbf{p}_t, x^h; \boldsymbol{\nu}^h) \equiv e(\mathbf{p}_v, v(\mathbf{p}_t, x^h; \boldsymbol{\nu}^h); \boldsymbol{\nu}^h).$$

[9] See for example, Lewbel (2001, equation 4), Hausman and Newey (2013, Assumption 1) and Blundell, Kristensen, and Matzkin (2014, Assumption A.1).

The money metric utility $\mu(\mathbf{p}_v, \mathbf{p}_t, x^h; \boldsymbol{\nu}^h)$ gives the minimal amount of expenditure that household h needs at prices $\mathbf{p}_v$ to be equally well of as it would have been when facing prices $\mathbf{p}_t$ and income x^h. The money metric utility lies at the basis of many cost of living indices. In particular, given two price vectors $\mathbf{p}_t$ and $\mathbf{p}_v$ and some reference budget $(\mathbf{p}, x)$, the Konüs cost of living index, describing the price increase from $\mathbf{p}_t$ to $\mathbf{p}_v$, is defined as,

$$\frac{\mu(\mathbf{p}_v, \mathbf{p}, x; \boldsymbol{\nu})}{\mu(\mathbf{p}_t, \mathbf{p}, x; \boldsymbol{\nu})}$$

There are two natural choices for $\mathbf{p}$, namely $\mathbf{p}_t$ and $\mathbf{p}_v$. Setting $\mathbf{p}$ equal to the initial price $\mathbf{p}_t$ gives the Laspeyres-Konüs cost of living index,

$$\frac{\mu(\mathbf{p}_v, \mathbf{p}_t, x; \boldsymbol{\nu})}{x}.$$

If we set $\mathbf{p}$ equal to the final price vector $\mathbf{p}_v$, we obtain the Paasche-Konüs cost of living index,

$$\frac{x}{\mu(\mathbf{p}_t, \mathbf{p}_v, x; \boldsymbol{\nu})},$$

Both indices are used to describe the increase in the cost necessary to maintain the same living standard over time. Their distributions can easily be constructed provided that we know the distribution of the money metric utility function. The money metric utility also provides a cardinalisation of the utility function in the sense that for any reference price vector $\mathbf{p}$ and for any two budgets $(\mathbf{p}_t, x)$ and $(\mathbf{p}_v, y)$:

$$\mu(\mathbf{p}, \mathbf{p}_t, x; \boldsymbol{\nu}^h) \geq \mu(\mathbf{p}, \mathbf{p}_v, y; \boldsymbol{\nu}^h) \iff v(\mathbf{p}_t, x; \boldsymbol{\nu}^h) \geq v(\mathbf{p}_v, y; \boldsymbol{\nu}^h).$$

As such, the difference in the money metric can be used as a measure for the welfare difference for two different budgets: if $(\mathbf{p}_t, x)$ is the old budget and $(\mathbf{p}_v, y)$ is the new one, then the welfare change can be measured by,

$$\mu(\mathbf{p}, \mathbf{p}_v, y; \boldsymbol{\nu}^h) - \mu(\mathbf{p}, \mathbf{p}_t, x; \boldsymbol{\nu}^h)$$

Again, there are two obvious choices for the base price vector $\mathbf{p}$, namely $\mathbf{p}_t$ or $\mathbf{p}_v$. The first leads to the equivalent variation,

$$EV = \mu(\mathbf{p}_t, \mathbf{p}_v, y; \boldsymbol{\nu}^h) - x.$$

The second gives the compensating variation,

$$CV = y - \mu(\mathbf{p}_v, \mathbf{p}_t, x; \boldsymbol{\nu}^h).$$

Revealed preferences The analysis in the following sections depends on a very simple revealed preference idea. Fix a household h and consider two distinct budgets $(\mathbf{p}_t, x)$ and $(\mathbf{p}_v, y)$. If the household is utility maximising, then the following condition must hold,

$$\text{If } x \geq \mathbf{p}_t\mathbf{q}(\mathbf{p}_v, y; \boldsymbol{\nu}^h) \text{ then } v(\mathbf{p}_t, x; \boldsymbol{\nu}^h) > v(\mathbf{p}_v, y; \boldsymbol{\nu}^h). \tag{5.1}$$

The reasoning behind the condition is simple, if $x \geq \mathbf{p}_t\mathbf{q}(\mathbf{p}_v, y; \boldsymbol{\nu}^h)$, then the consumed bundle $\mathbf{q}(\mathbf{p}_v, y; \boldsymbol{\nu}^h)$ at the budget $(\mathbf{p}_v, y)$ was also feasible when $\mathbf{q}(\mathbf{p}_t, x; \boldsymbol{\nu}^h)$ was chosen. Given that the household is utility maximising and that the budget sets are distinct, it follows that $u(\mathbf{q}(\mathbf{p}_t, x; \boldsymbol{\nu}^h); \boldsymbol{\nu}^h) > u(\mathbf{q}(\mathbf{p}_v, y; \boldsymbol{\nu}^h); \boldsymbol{\nu}^h)$, or equivalently, $v(\mathbf{p}_t, x; \boldsymbol{\nu}^h) > v(\mathbf{p}_v, y; \boldsymbol{\nu}^h)$. It is easy to see that Condition (5.1) implies the Weak Axiom of Revealed Preference (Samuelson (1938)), which states that for any two distinct budgets $(\mathbf{p}_t, x)$ and $(\mathbf{p}_v, y)$,

$$\text{If } x \geq \mathbf{p}_t\mathbf{q}(\mathbf{p}_v, y; \boldsymbol{\nu}^h) \text{ then } y < \mathbf{p}_v\mathbf{q}(\mathbf{p}_t, x; \boldsymbol{\nu}^h).$$

5.3 Nonparametric bounds

In this section we show how to use basic revealed preference restrictions, in particular Condition (5.1), together with information on the distribution of $\mathbf{q}(\mathbf{p}_t, x; \boldsymbol{\nu})$ in order to establish bounds on the distribution of the money metric utility function and the mean demand functions. As a first partial result, we demonstrate the possibility to obtain bounds on the proportion of households in the economy that prefer a certain budget over another.

Observational assumptions We depart from the observational restrictions imposed by a repeated cross sectional household consumption data set, where different households face the same prices in each cross section. This gives us a data structure that consists of a limited set of different price regimes, and for each price regime a large number of consumption bundles which are obtained from a random sample of households in the economy. We denote by $T = \{1, \ldots, |T|\}$, the set of cross sections. The price vector corresponding to cross section $t \in T$ is denoted by $\mathbf{p}_t$.

Given that different households face distinct expenditure levels, it is possible to observe (or estimate) the distribution of the random consumption bundles $\mathbf{q}(\mathbf{p}_t, x; \boldsymbol{\nu})$ for every cross sectional price vector $\mathbf{p}_t, (t \in T)$ and for any level of expenditure x. Actually, none of our results will require us to estimate the distributions of $\mathbf{q}(\mathbf{p}_t, x; \boldsymbol{\nu})$ but it will be easier to conceptualise things if we assume that these distributions are known. Estimation will be discussed in Section 5.4. We assume that $\mathbf{q}(\mathbf{p}_t, x; \boldsymbol{\nu})$ has a continuous density function which is strictly positive on its domain.

Before we start, let us introduce one last piece of notation. Let $A(\boldsymbol{\nu})$ represent a collection of conditions involving the random vector $\boldsymbol{\nu}$. We use the notation $\Pr(A(\boldsymbol{\nu}))$ as a shorthand for the following probability,

$$\Pr(A(\boldsymbol{\nu})) = \int \mathbb{1}[A(\boldsymbol{\nu})] dF(\boldsymbol{\nu}),$$

where $\mathbb{1}[\cdot]$ is the binary indicator function which equals one if and only if the term between brackets is true. $\Pr\left[A(\boldsymbol{\nu})\right]$ gives us the fraction of the households for which the statement $A(\boldsymbol{\nu}^h)$ is true. Equivalently, it gives us the probability that $A(\boldsymbol{\nu}^h)$ holds if we draw at random a household h from the population.

Using this notation, we further require that there is sufficient variation of preferences and demand such that for any two distinct budgets $(\mathbf{p}_t, x)$ and $(\mathbf{p}_v, y)$,

$$\Pr\left[x = \mathbf{p}_t \mathbf{q}(\mathbf{p}_v, y; \boldsymbol{\nu})\right] = 0, \text{ and}$$
$$\Pr\left[v(\mathbf{p}_t, x; \boldsymbol{\nu}) = v(\mathbf{p}_v, y; \boldsymbol{\nu})\right] = 0$$

This will allow us to freely interchange strict and weak inequalities within the function $\Pr[.]$.

5.3.1 Bounds on population preferences

In this section, we show how to establish bounds on the proportion of populations that prefer a certain budget $(\mathbf{p}_t, x)$ over another budget $(\mathbf{p}_v, y)$. Given the notation introduced above, this proportion is given by,

$$\Pr\left[v(\mathbf{p}_t, x; \boldsymbol{\nu}) \geq v(\mathbf{p}_v, y; \boldsymbol{\nu})\right].$$

The following shows how to obtain bounds on this proportion using only information on the distribution of $\mathbf{q}(\mathbf{p}_v, y; \boldsymbol{\nu})$. Consider the fraction of households for which $x \geq \mathbf{p}_t \mathbf{q}(\mathbf{p}_v, y; \boldsymbol{\nu}^h)$.

$$r_{t,v}(x, y) = \Pr\left[x \geq \mathbf{p}_t \mathbf{q}(\mathbf{p}_v, y; \boldsymbol{\nu})\right]$$

We claim that this number provides a lower bound on the fraction of households that prefer the budget $(\mathbf{p}_t, x)$ over the budget $(\mathbf{p}_v, y)$.

Lemma 5.2.

$$r_{t,v}(x,y) \leq \Pr\left[v(\mathbf{p}_t, x; \boldsymbol{\nu}) \geq v(\mathbf{p}_v, y; \boldsymbol{\nu})\right].$$

Proof. Given that all households are rational, we know from Condition (5.1) that for all values $\boldsymbol{\nu}^h$ of the random vector:

$$\text{If } x \geq \mathbf{p}_t \mathbf{q}(\mathbf{p}_v, y; \boldsymbol{\nu}^h) \text{ then, } v(\mathbf{p}_t, x; \boldsymbol{\nu}^h) \geq v(\mathbf{p}_v, y; \boldsymbol{\nu}^h).$$

This means that $\mathbb{1}\left[x \geq \mathbf{p}_t \mathbf{q}(\mathbf{p}_v, y; \boldsymbol{\nu}^h)\right] \leq \mathbb{1}\left[v(\mathbf{p}_t, x; \boldsymbol{\nu}^h) \geq v(\mathbf{p}_v, y; \boldsymbol{\nu}^h)\right]$. Integrating both sides over all values of the random vector $\boldsymbol{\nu}$ and using Assumption 5.1, we obtain,

$$\begin{aligned} &\Pr\left[x \geq \mathbf{p}_t \mathbf{q}(\mathbf{p}_v, y; \boldsymbol{\nu})\right] \leq \Pr\left[v(\mathbf{p}_t, x; \boldsymbol{\nu}) \geq v(\mathbf{p}_v, y; \boldsymbol{\nu})\right], \\ \iff &r_{t,v}(x,y) \leq \Pr\left[v(\mathbf{p}_t, x; \boldsymbol{\nu}) \geq v(\mathbf{p}_v, y; \boldsymbol{\nu})\right]. \end{aligned}$$

□

Figure 5.1 illustrates the reasoning behind the lemma in the two goods case. The figure gives two budget sets $(\mathbf{p}_v, y)$ and $(\mathbf{p}_t, x)$. Given budget exhaustion, consumption within each budget set is distributed over the respective budget lines. All consumption bundles on the dashed segment capture the consumption bundles of the households that satisfy $x > \mathbf{p}_t \mathbf{q}(\mathbf{p}_v, y; \boldsymbol{\nu}^h)$. As such, the mass of households that consume on this line segment is equal to $r_{t,v}(x,y)$. What the lemma says is that this fraction is smaller than the proportion of households for which $v(\mathbf{p}_t, x; \boldsymbol{\nu}^h) \geq v(\mathbf{p}_v, y; \boldsymbol{\nu}^h)$. If this would not be the case, then there would be an individual with $v(\mathbf{p}_t, x; \boldsymbol{\nu}^h) < v(\mathbf{p}_v, y; \boldsymbol{\nu}^h)$ who consumes a bundle on the dashed line segment when facing the budget $(\mathbf{p}_v, y)$. However, this is impossible because this consumption bundle is in the interior of the budget set $(\mathbf{p}_t, x)$, which means, using Condition (5.1), that $v(\mathbf{p}_t, x; \boldsymbol{\nu}^h) > v(\mathbf{p}_v, y; \boldsymbol{\nu}^h)$.

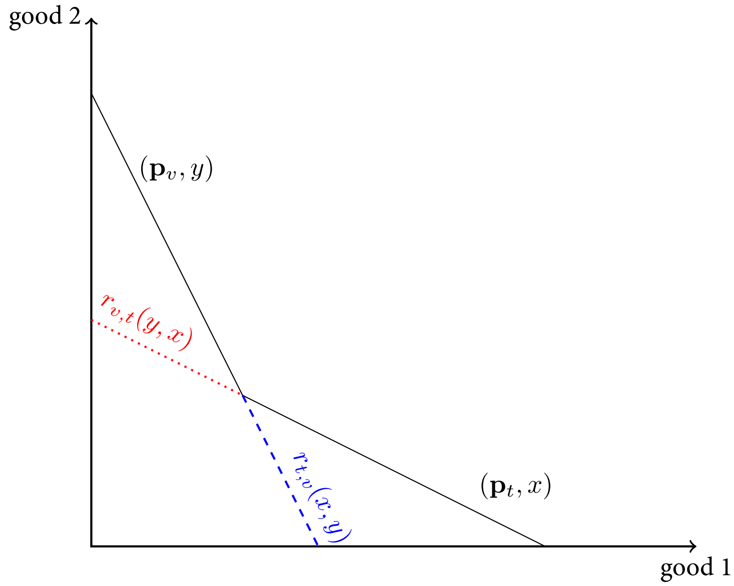


Figure 5.1: Illustration of Lemma 5.2

Given above lemma, and the fact that,

$$\Pr\left[v(\mathbf{p}_t, x; \boldsymbol{\nu}) \geq v(\mathbf{p}_v, y; \boldsymbol{\nu})\right] + \Pr\left[v(\mathbf{p}_v, y; \boldsymbol{\nu}) \geq v(\mathbf{p}_t, x; \boldsymbol{\nu})\right] = 1,$$

We immediately obtain the upper bound,

$$\Pr\left[v(\mathbf{p}_t, x; \boldsymbol{\nu}) \geq v(\mathbf{p}_v, y; \boldsymbol{\nu})\right] \leq 1 - r_{v,t}(y, x).$$

For both lower and upper bounds to be valid, it should be the case that for all cross sections $t, v \in T$ and all expenditure levels x, y,

$$r_{t,v}(x, y) + r_{v,t}(y, x) \leq 1.$$

This condition is equivalent to the Weak Axiom of Stochastic Revealed Preference applied to our setting (see Bandyopadhyay, Dasgupta, and Pattanaik (1999, 2002, 2004)). Hoderlein and Stoye (2014) and Kawaguchi (2012) recently developed (among other things) a statistical test that verifies (for two given budgets $(\mathbf{p}_t, x)$ and $(\mathbf{p}_v, y)$) whether this condition is satisfied. If we go back to Figure 5.1, the condition states that the sum of the mass of households on the dashed line segment and the mass of households on the dotted line segment must be smaller than 1. If this would not be the case, then there would be a household which is on the dashed segment when the budget is $(\mathbf{p}_v, y)$ and on the dotted segment when the budget is equal to $(\mathbf{p}_t, x)$. However, this implies that the household violates the Weak Axiom of Revealed Preference (i.e. $x \geq \mathbf{p}_t \mathbf{q}(\mathbf{p}_v, y; \boldsymbol{\nu}^h)$ and $y \geq \mathbf{p}_v \mathbf{q}(\mathbf{p}_t, x; \boldsymbol{\nu}^h)$).

The condition also shows that our bounds on $\Pr\left[v(\mathbf{p}_t, x; \boldsymbol{\nu}) \geq v(\mathbf{p}_v, y; \boldsymbol{\nu})\right]$ will be tighter, the closer $r_{t,v}(x, y) + r_{v,t}(y, x)$ is to one. In particular, if the sum equals one, then $\Pr\left[v(\mathbf{p}_t, x; \boldsymbol{\nu}) \geq v(\mathbf{p}_v, y; \boldsymbol{\nu})\right]$ will be exactly identified.

There are two potential issues that may arise. First of all, it may happen that $r_{t,v}(x, y) + r_{v,t}(y, x)$ is larger than one, in which case the bounds cannot be simultaneously satisfied. Alternatively, it may happen that $r_{t,v}(x, y) + r_{v,t}(y, x)$ is considerably smaller than one, in which case the range may be too large to contain much useful information. Whether one of those problems arises is obviously an empirical matter. However, it may nevertheless be useful to discuss each of the issues a bit more in detail and to present some potential solutions.

Incompatible bounds A first problem arises if,

$$r_{t,v}(x, y) + r_{v,t}(y, x) > 1,$$

for some budgets $(\mathbf{p}_t, x)$ and $(\mathbf{p}_v, y)$. In such case, we know that there is at least one household in the population that violates the Weak Axiom of Revealed Preference, hence, we should reject rationality of all households in the population. In order to remedy the situation, we see two possible solutions. A first solution is to allow a certain fraction of the population to violate the Weak Axiom of Revealed Preference, i.e. a certain subset of the population is considered to be irrational. Applying this solution would amount to subtracting a certain percentage, that equals the fraction of irrational households, from $r_{t,v}(x,y)$ and $r_{v,t}(y,x)$, thereby widening the range of possible values for $\Pr\left[v(\mathbf{p}_t, x; \boldsymbol{\nu}) \geq v(\mathbf{p}_v, y; \boldsymbol{\nu})\right]$.

A second solution is to relax the rationality constraints for all households simultaneously. In revealed preference theory, such relaxation is usually conceived by using a 'goodness–of–fit' measure. The basic idea here is that households may not 'exactly' pass the revealed preference restrictions but are still very close to passing them. The way to proceed is to consider an extended version of the basic revealed preference conditions that focuses on nearly optimising behaviour rather than exactly optimising behaviour. See also Varian (1990) for a general discussion on the usefulness of considering such nearly optimising behaviour in empirical revealed preference analysis. Here, we consider one way in which this can be accomplished. We consider an adaptation of an early proposal of Afriat (1973) for revealed preference tests in a non-stochastic setting to our specific setting. In particular, we capture optimisation error by a so-called Afriat index $e \in [0,1]$. For a given value of e, the new rationality criterion adjusts condition (5.1) in the following way,

$$\text{If } e \cdot x \geq \mathbf{p}_t \mathbf{q}(\mathbf{p}_v, y; \boldsymbol{\nu}^h) \text{ then } v(\mathbf{p}_t, x; \boldsymbol{\nu}^h) > v(\mathbf{p}_v, y; \boldsymbol{\nu}^h).$$

When comparing the budget $(\mathbf{p}_t, x)$ to another budget $(\mathbf{p}_v, y)$ the Afriat index e reduces the expenditure level x towards $e \cdot x$. In other words, we now check whether behaviour is rational while allowing the household to waste as much as $(1-e)$ of the income x by making irrational choices. In other words, we only require the household to prefer the bundle $\mathbf{q}(\mathbf{p}_t, x, \boldsymbol{\nu})$ over the bundle $\mathbf{q}(\mathbf{p}_v, y, \boldsymbol{\nu})$ if the latter is in the budget $(\mathbf{p}_t, e \cdot x)$. As such, wasting/irrational behaviour can also be regarded as sub-optimising behaviour, we thus verify whether behaviour is rational if we account for an optimisation error equal to $(1-e)$.

Lowering the value of e will lead to a less strict test. Using this Afriat index, we can construct the following probabilities,

$$r_t^e(x, y) = \Pr\left[e \cdot x \geq \mathbf{p}_t \mathbf{q}(\mathbf{p}_v, y; \boldsymbol{\nu})\right].$$

The number $r_t^e(x, y)$ is increasing in e and $r_{t,v}^0(x, y) = 0$. Given this, there will always be a value of $e \in [0, 1]$ such that,

$$r_{t,v}^e(x, y) + r_{v,t}^e(y, x) \leq 1.$$

Notice that when $e \cdot x \geq \mathbf{p}_t \mathbf{q}(\mathbf{p}_v, y; \boldsymbol{\nu}^h)$ and $e \cdot y \geq \mathbf{p}_v \mathbf{q}(\mathbf{p}_t, x; \boldsymbol{\nu}^h)$, we have that $v(\mathbf{p}_t, x; \boldsymbol{\nu}^h) > v(\mathbf{p}_v, y; \boldsymbol{\nu}^h)$ and $v(\mathbf{p}_t, x; \boldsymbol{\nu}^h) < v(\mathbf{p}_v, y; \boldsymbol{\nu}^h)$, an impossibility. Indeed, given a fixed level of e, it is still possible to reject rationality. The analysis could then proceed by replacing $r_t(x, y)$ by the numbers $r_t^{e^*}(x, y)$ where e^* is the largest number for which the above inequalities are consistent.

Uninformative bounds A second problem arises if $r_{t,v}(x, y) + r_{v,t}(y, x)$ is considerably below 1. In such cases, the range of values for $\Pr\left[v(\mathbf{p}_t, x; \boldsymbol{\nu}) > v(\mathbf{p}_v, y; \boldsymbol{\nu})\right]$ will be too wide and, therefore, not very informative. An approach to tighten the bounds is to impose a stronger stochastic revealed preference condition. In the construction of $r_{t,v}(x, y)$ above, we only used information concerning the two budget sets $(\mathbf{p}_t, x)$ and $(\mathbf{p}_v, y)$. In some cases, however, it is possible to include information on additional budget sets to obtain tighter bounds. One such tightening relies on the fact that for any three distinct numbers a, b and c it is always the case that,

$$\Pr(a > c) \geq \Pr(a > b) + \Pr(b > c) - 1.$$

The reasoning behind the inequality is simple. The probability that c is larger than b is given by $1 - \Pr(b > c)$. As such, $\Pr(a > b > c)$ is bounded from below by $\Pr(a > b) - (1 - \Pr(b > c))$. The conclusion then follows from the fact that $\Pr(a > c) \geq \Pr(a > b > c)$. Rewriting above condition shows that it is equivalent to the famous triangle inequality.

$$\Pr(b > c) \leq \Pr(b > a) + \Pr(a > c).$$

The triangle inequality has first been noted by Guilbaud (1953) and has been popularised by Marschak (1960). The inequality is one of the key conditions in the literature on binary probability systems. This literature, which is closely related to the literature on stochastic revealed preference theory, tries to characterise all collections of binary probabilities over a finite set of alternatives that are induced by probability distributions over the family of linear orders (preference relations) on this set.

If we apply above condition to our setting and use the previously established bounds, we obtain that for all $t, v, w \in T$ and all incomes x, y, z,

$$\begin{aligned}
&\Pr\left[v(\mathbf{p}_t, x; \boldsymbol{\nu}) \geq v(\mathbf{p}_v, y; \boldsymbol{\nu})\right] \\
&\geq \Pr\left[v(\mathbf{p}_t, x; \boldsymbol{\nu}) \geq v(\mathbf{p}_w, z; \boldsymbol{\nu})\right] \\
&+ \Pr\left[v(\mathbf{p}_w, z; \boldsymbol{\nu}) \geq v(\mathbf{p}_v, y; \boldsymbol{\nu})\right] - 1, \\
&\geq r_{t,w}(x, z) + r_{w,v}(z, y) - 1.
\end{aligned}$$

In cases where $r_{t,v}(x, y)$ is lower than

$$\max_{w,z} \left\{r_{t,w}(x, z) + r_{w,v}(z, y) - 1\right\},$$

this improves the lower bound on $\Pr[v(\mathbf{p}_t, x; \boldsymbol{\nu}) \geq v(\mathbf{p}_v, y; \boldsymbol{\nu})]$. Of course this tightening of the bounds can be iterated until no further improvements are possible. If the range is still too wide, further tightening could still be obtained by using other, though more elaborate 'binary probability system' conditions. See, for example, Fishburn (1992) for an overview of the various kinds of conditions that could be imposed. Although the triangle inequality potentially improves the bounds, we found that for our application, it does not give any significant improvements. The main reason is probably that, in our application, the bounds are already quite narrow. Given this, we abstain from implementing it in the empirical analysis.

5.3.2 Bounds on money metric utility

In this section, we show how to use above results to bound the distribution of the money metric utility function $\mu(\mathbf{p}_t, \mathbf{p}_0, x_0; \nu)$ for some price vectors $\mathbf{p}_0$ and $\mathbf{p}_t$ corresponding to the prices of two cross sections in the data set and for a particular level of income x_0. Let us first focus on the upper bounds.

Upper bounds Let us fix a cross sectional price vector $\mathbf{p}_0$ and an income level x_0. For any number $\pi \in (0, 1)$ and any cross section $t \in T$, let $h_t(\pi)$ solve the following condition,

$$\begin{aligned}\pi &= \Pr\left[h_t(\pi) \geq \mathbf{p}_t\mathbf{q}(\mathbf{p}_0, x_0; \boldsymbol{\nu})\right],\\ &= r_{t,0}(h_t(\pi), x_0)\end{aligned}$$

The value of $h_t(\pi)$ corresponds to the πth quantile of the random variable $\mathbf{p}_t\mathbf{q}(\mathbf{p}_0, x_0; \boldsymbol{\nu})$. From Lemma 5.2, we know that π is lower than the fraction of the households that prefer the budget $(\mathbf{p}_t, h_t(\pi))$ over the budget $(\mathbf{p}_0, x_0)$.

$$\begin{aligned}\pi &\leq \Pr\left[v(\mathbf{p}_t, h_t(\pi); \boldsymbol{\nu}) \geq v(\mathbf{p}_0, x_0; \boldsymbol{\nu})\right],\\ &= \Pr\left[h_t(\pi) \geq \mu(\mathbf{p}_t, \mathbf{p}_0, x_0; \boldsymbol{\nu})\right].\end{aligned}$$

The second line is obtained by inverting the indirect utility function $v(\mathbf{p}_t, h_t(\pi); \boldsymbol{\nu})$ with respect to its second argument. This can be done by the fact that the indirect utility function is strictly increasing in income.

Let us denote by $m_t(\pi)$ the quantile function of $\mu(\mathbf{p}_t, \mathbf{p}_0, x_0)$, i.e. for all $\pi \in (0, 1)$

$$\Pr\left[\mu(\mathbf{p}_t, \mathbf{p}_0, x_0; \boldsymbol{\nu}) \leq m_t(\pi)\right] = \pi.$$

Then, using our previously established result, we have that,

$$\begin{aligned} &\Pr\left[\mu(\mathbf{p}_t, \mathbf{p}_0, x_0; \boldsymbol{\nu}) \leq m_t(\pi)\right] = \pi \leq \Pr\left[\mu(\mathbf{p}_t, \mathbf{p}_0, x_0; \boldsymbol{\nu}) \leq h_t(\pi)\right], \\ \iff &m_t(\pi) \leq h_t(\pi). \end{aligned}$$

The last line uses the assumption that the cumulative distribution function of $\mu(\mathbf{p}_t, \mathbf{p}_0, x; \boldsymbol{\nu})$ is strictly increasing on its support. This result shows that $h_t(\pi)$ is an upper bound on the πth quantile of the distribution of the money metric utility function. Using these upper bounds on the quantiles; we can also derive an upper bound on the mean value of the money metric utility. Let M be the mean of the function $\mu(\mathbf{p}_t, \mathbf{p}_0, x_0; \boldsymbol{\nu})$. We have that:

$$\begin{aligned} M &= \int_0^\infty \mu(\mathbf{p}_t, \mathbf{p}_0, x_0; \boldsymbol{\nu}) dF(\mu(\mathbf{p}_t, \mathbf{p}_0, x_0, \boldsymbol{\nu})), \\ &= \int_0^1 m_t(\pi) d\pi \leq \int_0^1 h_t(\pi) d\pi. \end{aligned}$$

In practice, we compute the values of $h_t(\pi)$ for a finite grid of values $\pi_0, \pi_1, \ldots, \pi_n$ with $\pi_0 = 0$ and $\pi_N = 1$.[10] This allows us to approximate this upper bound by,

$$\int_0^1 h_t(\pi) d\pi \leq \sum_{n=1}^{N} (\pi_n - \pi_{n-1}) h_t(\pi_n).$$

[10] The upper bound $h_t(1)$ can be set to the minimal income such that the budget hyperplane for $(\mathbf{p}_t, h_t(1))$ lies above the hyperplane for $(\mathbf{p}_0, x_0)$.

The finer the grid, the better the approximation.

Lower bounds We use a similar procedure to compute lower bounds on the quantiles. For $\pi \in (0,1)$, let $\ell_t(\pi)$ solve the following equality,

$$1 - \pi = \Pr\left[x_0 \geq \mathbf{p}_0 \mathbf{q}(\mathbf{p}_t, \ell_t(\pi); \boldsymbol{\nu})\right],$$
$$= r_{0,t}(x_0, \ell_t(\pi))$$

Then,

$$1 - \pi \leq \Pr\left[v(\mathbf{p}_0, x_0; \boldsymbol{\nu}) \geq v(\mathbf{p}_t, \ell_t(\pi); \boldsymbol{\nu})\right],$$
$$= \Pr\left[\mu(\mathbf{p}_t, \mathbf{p}_0, x_0; \boldsymbol{\nu}) \geq \ell_t(\pi)\right],$$
$$= 1 - \Pr\left[\mu(\mathbf{p}_t, \mathbf{p}_0, x_0; \boldsymbol{\nu}) \leq \ell_t(\pi)\right]$$

As before, let $m_t(\pi)$ be the πth quantile of the distribution of the money metric utility $\mu(\mathbf{p}_t, \mathbf{p}_0, x_0; \boldsymbol{\nu})$. We have that,

$$\Pr\left[\mu(\mathbf{p}_t, \mathbf{p}_0, x_0; \boldsymbol{\nu}) \leq m_t(\pi)\right] = \pi \geq \Pr\left[\mu(\mathbf{p}_t, \mathbf{p}_0, x_0; \boldsymbol{\nu}) \leq \ell_t(\pi)\right],$$
$$\iff m_t(\pi) \geq \ell_t(\pi)$$

This shows that $\ell_t(\pi)$ is a lower bound for the quantile $m_t(\pi)$. The mean M is then bounded from below by the quantity $\int_0^1 \ell_t(\pi) d\pi$ which can be approximated by $\sum_{n=0}^{N-1} \ell_t(\pi_n)(\pi_{n+1} - \pi_n)$.[11]

[11] The lower bound $\ell_t(0)$ can be set to the maximal income such that the hyperplane for the budget set $(\mathbf{p}_t, \ell_t(0))$ lies below the hyperplane for $(\mathbf{p}_0, x_0)$.

5.3.3 Bounds on demand correspondences

In this section, we show how to adapt above framework in order to establish bounds on the quantiles of the demand functions for unobserved, counterfactual, price regimes $\mathbf{p}_0$ and expenditure levels x_0, i.e. $\mathbf{p}_0$ does not necessarily correspond to a price vector of a certain cross section.

Consider a function $f : \mathbb{R}^n_+ \to \mathbb{R} : \mathbf{q} \mapsto f(\mathbf{q})$. In this section, we will provide upper bounds on the quantiles of the distribution of the random variable $f(\mathbf{q}(\mathbf{p}_0, x_0, \boldsymbol{\nu}))$. The function $f(.)$ encompasses various interesting measures. For example, if we want to bound the expenditure share on one of the goods, we can use the function $f(\mathbf{q}) = \frac{1}{x_0} p_{0,j} q_j$, where $p_{0,j}$ is the price of good j in vector $\mathbf{p}_0$, q_j is the quantity of good j in vector $\mathbf{q}$ and x_0 is the expenditure level.

The focus on upper bounds is not restrictive given that we can always use information on upper bounds to construct lower bounds. In order to see this, let $-m(1-\pi)$ be the $(1-\pi)$th quantile of the variable $-f(\mathbf{q}(\mathbf{p}_0, x_0; \boldsymbol{\nu}))$ and let $-g(1-\pi)$ be its upper bound. We then have that,

$$
\begin{aligned}
1 - \pi &= \int \mathbb{1}\left[-f(\mathbf{q}(\mathbf{p}_0, x_0, \boldsymbol{\nu})) \leq -m(1-\pi)\right] dF(\boldsymbol{\nu}),\\
&\leq \int \mathbb{1}\left[-f(\mathbf{q}(\mathbf{p}_0, x_0, \boldsymbol{\nu})) \leq -g(1-\pi)\right] dF(\boldsymbol{\nu}),\\
\iff \qquad \pi &\geq 1 - \int \mathbb{1}\left[-f(\mathbf{q}(\mathbf{p}_0, x_0, \boldsymbol{\nu})) \leq -g(1-\pi)\right] dF(\boldsymbol{\nu}),\\
&= \int \mathbb{1}\left[-f(\mathbf{q}(\mathbf{p}_0, x_0, \boldsymbol{\nu})) > -g(1-\pi)\right] dF(\boldsymbol{\nu}),\\
&= \int \mathbb{1}\left[f(\mathbf{q}(\mathbf{p}_0, x_0, \boldsymbol{\nu})) \leq g(1-\pi)\right] dF(\boldsymbol{\nu}).
\end{aligned}
$$

As such, we see that $g(1-\pi)$ provides a lower bound on the πth quantile of $f(\mathbf{q}(\mathbf{p}_0, x_0; \boldsymbol{\nu}))$. For example, we can establish a lower bound on the πth quantile of $f(\mathbf{q}) = \frac{1}{x_0} p_{0,j} q_j$ by constructing an upper bound on the $(1-\pi)$th quantile of $-\frac{1}{x_0} p_{0,j} q_j (= \sum_{i \neq j} \frac{1}{x_0} p_{0,i} q_i - 1)$.

For every cross section t, we previously defined the value $\ell_t(1-\pi)$ that satisfied the following condition,

$$\begin{aligned}\pi &= \Pr\left[x_0 \geq \mathbf{p}_0\mathbf{q}(\mathbf{p}_t, \ell_t(1-\pi); \boldsymbol{\nu})\right],\\ &= r_{0,t}(x_0, \ell_t(1-\pi)).\end{aligned}$$

The value of $\ell_t(1-\pi)$ can be obtained using information on x_0, $\mathbf{p}_0$ and the distribution of $\mathbf{q}(\mathbf{p}_t, x; \boldsymbol{\nu})$ alone, which we assumed to be known. For the next step, we use the Weak Axiom of Stochastic Revealed Preference, which requires that,

$$\begin{aligned}& r_{t,0}(\ell_t(1-\pi), x_0) + r_{0,t}(x_0, \ell_t(1-\pi)) \leq 1,\\ \Longleftrightarrow\ & r_{0,t}(x_0, \ell_t(1-\pi)) \leq 1 - r_{t,0}(\ell_t(1-\pi), x_0).\end{aligned}$$

Let $m(\pi)$ be the πth quantile of the distribution function of the random variable $f(\mathbf{q}(\mathbf{p}_0, x_0, \boldsymbol{\nu}))$. We have that,

$$\begin{aligned}\Pr\left[f(\mathbf{q}(\mathbf{p}_0, x_0; \boldsymbol{\nu}) \leq m(\pi)\right] = \pi &= r_{0,t}(x_0, \ell_t(1-\pi)),\\ &\leq 1 - r_{t,0}(\ell_t(1-\pi), x_0),\\ &= \Pr\left[\ell_t(1-\pi) \leq \mathbf{p}_t\mathbf{q}(\mathbf{p}_0, x_0; \boldsymbol{\nu})\right]\\ &\leq \Pr\left[\begin{array}{c} f(\mathbf{q}(\mathbf{p}_0, x_0, \boldsymbol{\nu})) \leq \max_{\mathbf{q}} f(\mathbf{q}) \\ \text{s.t } \ell_t(1-\pi) \leq \mathbf{p}_t\mathbf{q} \text{ and } \mathbf{p}_0\mathbf{q} = x_0 \end{array}\right]\end{aligned}$$

The last inequality follows from the fact that whenever $\ell_t(1-\pi) \leq \mathbf{p}_t\mathbf{q}(\mathbf{p}_0, x_0; \boldsymbol{\nu})$ holds, then $f(\mathbf{q}(\mathbf{p}_0, x_0, \boldsymbol{\nu})) \leq \max_{\mathbf{q}} f(\mathbf{q})$ s.t $\ell_t(1-\pi) \leq \mathbf{p}_t\mathbf{q}$ and $\mathbf{p}_0\mathbf{q} = x_0$ must also hold. In order to see this, assume on the contrary that $f(\mathbf{q}(\mathbf{p}_0, x_0, \boldsymbol{\nu}))$ is larger than $f(\mathbf{q})$ for all vectors $\mathbf{q}$ where $\mathbf{p}_0\mathbf{q} = x_0$ and $\ell_t(1-\pi) \leq \mathbf{p}_t\mathbf{q}$. Then, given that $\mathbf{p}_0\mathbf{q}(\mathbf{p}_0, x_0; \boldsymbol{\nu}) = x_0$, it must be that $\ell_t(1-\pi) > \mathbf{p}_t\mathbf{q}(\mathbf{p}_0, x_0; \boldsymbol{\nu})$, a contradiction.

Above result shows that,

$$m(\pi) \leq \max_{\mathbf{q}} f(\mathbf{q}) \text{ s.t } \ell_t(1-\pi) \leq \mathbf{p}_t\mathbf{q} \text{ and } \mathbf{p}_0\mathbf{q} = x_0,$$

for all cross sections t. In practice, we compute this right hand side for every cross section t and then take the lowest value across all cross sections as the upper bound. If f is a linear function, then the right hand side is a simple linear programming problem which can be solved efficiently.

The construction of the bounds in the simple two goods setting is illustrated in Figure 5.2. There are three budget lines corresponding to $(\mathbf{p}_0, x_0)$, $(\mathbf{p}_t, \ell_t(1-\pi))$ and $(\mathbf{p}_v, \ell_v(\pi))$. The incomes $\ell_t(1-\pi)$ and $\ell_v(\pi)$ are chosen such that the mass of households on the dashed line segment (where $x_0 > \mathbf{p}_0\mathbf{q}(\mathbf{p}_t, \ell_t(1-\pi), \boldsymbol{\nu})$) is equal to π and the mass of households on the dash-dotted line segment (where $x_0 > \mathbf{p}_0\mathbf{q}(\mathbf{p}_v, \ell_v(\pi), \boldsymbol{\nu})$) is equal to $(1-\pi)$.

The quantity $\overline{q}_2$ is the maximum value of good 2 that corresponds to a bundle on the budget $(\mathbf{p}_0, x_0)$ (where $\mathbf{p}_0\mathbf{q} = x_0$) and $\ell_t(1-\pi) \leq \mathbf{p}_t\mathbf{q}$. From the result above, we know that this value gives an upper bound on the πth quantile of the distribution of $q_2(\mathbf{p}_0, x_0; \boldsymbol{\nu})$. Given that there are only two goods, this upper bound immediately gives a lower bound on the $(1-\pi)$th quantile of $q_1(\mathbf{p}_0, x_0; \boldsymbol{\nu})$, given by $\underline{q}_1$. Similarly, $\overline{q}_1$ gives an upper bound on the $(1-\pi)$th quantile of $q_1(\mathbf{p}_0, x_0; \boldsymbol{\nu})$, while $\underline{q}_2$ gives a lower bound on the πth quantile of $q_2(\mathbf{p}_0, x_0; \boldsymbol{\nu})$. As such, the πth quantile of $q_2(\mathbf{p}_0, x_0; \boldsymbol{\nu})$ is bounded by the quantities $\overline{q}_2$ and $\underline{q}_2$. Given these bounds on the quantiles of the demand functions, we can compute bounds on the mean of the demand function by using a similar procedure as for the money metric utility function.

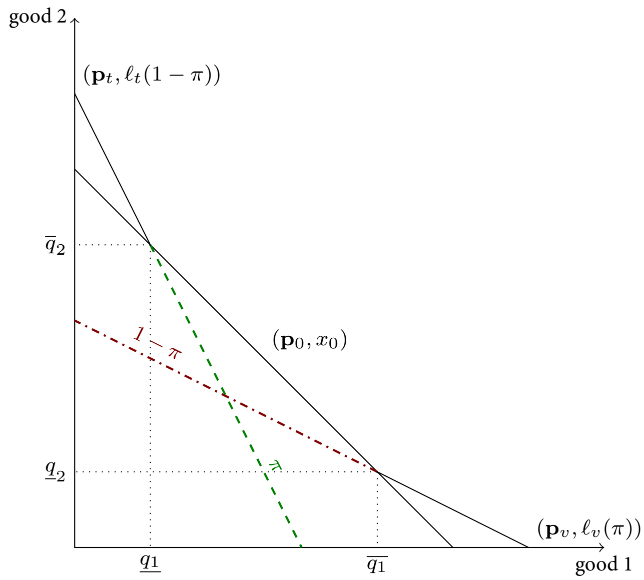


Figure 5.2: Illustration of the construction of the bounds

5.4 Application

In this section, we discuss the empirical implementation of the theoretical bounds that were established in the previous section. We first present our estimation procedure for the measures $r_{t,v}(x, y)$, $\ell_t(\pi)$ and $h_t(\pi)$. Next, we discuss a modification of the estimator with better finite sample properties and we show how we control for observed heterogeneity and endogeneity of the total expenditures. We also very briefly discuss the issue of statistical inference on bounds. Finally, we present some empirical results.

5.4.1 Estimation procedure

The construction of the bounds in the previous section assumed that we know the distribution of the variables $\mathbf{q}(\mathbf{p}_t, x; \boldsymbol{\nu})$ for every cross sectional price $\mathbf{p}_t$ and every income level x. Given these distributions it is fairly straightforward to obtain the quantities $r_{t,v}(x, y) = \Pr[x \geq \mathbf{p}_t\mathbf{q}(\mathbf{p}_v, y; \boldsymbol{\nu})]$, which form the main building blocks for our bounds. In practice, however, these probabilities need to be estimated. We propose a kernel estimator. For the estimation of the numbers $\ell_t(\pi)$ and $h_t(\pi)$ which are used to bound the money metric utility functions and the demand functions we propose a plug-in estimator.

Consider the vth cross section, $v \in T$. Assume that this cross section contains a sample of n observed household demand bundles $\{\mathbf{q}_{v,i}\}_{i\leq n}$ where i corresponds to a particular observation. We denote by $\{x_{v,i}\}_{i\leq n}$ the corresponding expenditure levels ($x_{v,i} = \mathbf{p}_v\mathbf{q}_{v,i}$). We assume that the sample $\{\mathbf{q}_{v,i}\}_{i\leq n}$ is i.i.d drawn from the random vector $\mathbf{q}_v$. We denote by x_v the random variable $\mathbf{p}_v\mathbf{q}_v$. We denote by $\varphi(.)$ the distribution function of x_v.

Consider the value $r_{t,v}(x,y) = \Pr\left[x \geq \mathbf{p}_t\mathbf{q}(\mathbf{p}_v, y; \boldsymbol{\nu})\right]$. This value corresponds to the cdf of the random variable $\mathbf{p}_t\mathbf{q}_v$, conditional on the value $x_v = y$,

$$\begin{aligned} r_{t,v}(x,y) &= \int \mathbb{1}\left[x \geq \mathbf{p}_t\mathbf{q}(\mathbf{p}_v, y, \boldsymbol{\nu})\right] dF(\boldsymbol{\nu}), \\ &= \int \mathbb{1}\left[x \geq \mathbf{p}_t\mathbf{q}_v\right] dF(\mathbf{q}_v|x_v = y), \end{aligned}$$

where $F(\mathbf{q}_v|x_v = y)$ is the conditional cdf of $\mathbf{q}_v$ given the level of expenditure $x_v = y$. This expression is equal to the conditional mean of the indicator function $\mathbb{1}\left[x \geq \mathbf{p}_t\mathbf{q}_v\right]$,

$$\int \mathbb{1}\left[x \geq \mathbf{p}_t\mathbf{q}_v\right] dF(\mathbf{q}_v|x_v = y) = \mathbb{E}\left\{\mathbb{1}\left[x - \mathbf{p}_t\mathbf{q}_v \geq 0\right]|x_v = y\right\} \equiv g(\ln(y)).$$

We can express this conditional mean as,

$$\mathbb{1}\left[x - \mathbf{p}_t\mathbf{q}_{v,i} \geq 0\right] = g(\ln(x_{v,i})) + \varepsilon_{v,i},$$

where $\mathbb{E}(\varepsilon_{v,i}|x_v = y) = 0$ for all values of y. The quantity of interest is given by the value of $g(\ln(y))$. A straightforward Nadaraya-Watson kernel estimator is given by,

$$\hat{r}_{t,v}(x,y) = \frac{\frac{1}{nh}\sum_{i=1}^{n} \mathbb{1}\left[x \geq \mathbf{p}_t \mathbf{q}_{v,i}\right] k\left(\frac{\ln(x_{v,i})-\ln(y)}{h}\right)}{\frac{1}{nh}\sum_{i=1}^{n} k\left(\frac{\ln(x_{v,i})-\ln(y)}{h}\right)}.$$

where h is the bandwidth and $k\,(.)$ is a symmetric kernel function that satisfies $\int k(v)dv = 1$ and $\int vk(v)dv = 0$.[12] We take the log of expenditure as metric.

If for $n \to \infty$, $h \to 0$ and $nh \to \infty$, then the estimator $\hat{r}_{t,v}(x,y)$ consistently estimates $r_{t,v}(x,y)$. If in addition (i) $\varphi(y) > 0$, (ii) $r_{t,v}(x,y) \in (0,1)$, (iii) the functions $\varphi(.)$ and $r_{t,v}(.,.)$ are sufficiently smooth[13] and (iv) $nh^7 \to 0$ then the estimator $\hat{r}_{t,v}(x,y)$ has the following asymptotic distribution (see, for example Li and Racine (2007)).

$$\sqrt{nh}\frac{1}{V(x,y)^{1/2}}\left[\hat{r}_{t,v}(x,y) - r_{t,v}(x,y) - B(x,y)\right] \to N(0,1).$$

where

$$B(x,y) = h^2\kappa_2\left[\frac{1}{2}\frac{\partial^2 r_{t,v}(x,y)}{\partial y^2} + \frac{\partial r_{t,v}(x,y)}{\partial y}\frac{\partial\varphi(y)}{\partial y}\frac{1}{\varphi(y)}\right]$$

is the asymptotic bias and

$$V(x,y) = r_{t,v}(x,y)\left[1 - r_{t,v}(x,y)\right]\kappa/\varphi(y)$$

is the asymptotic variance. Here $\kappa_2 = \int v^2 k(v)dv$ and $\kappa = \int k(v)^2 dv$. As usual with nonparametric kernel estimators, the bias, $B(y,x)$, does not disappear asymptotically when using the optimal bandwidth $h = O(n^{1/5})$. One possible solution is to undersmooth.

[12] In practice, we use the Gaussian kernel.

[13] The exact condition is that $\varphi(y)$ and $r_{t,v}(x,y)$ have continuous second order derivatives with respect to y.

The estimates for $h_t(\pi)$ and $\ell_t(\pi)$ are computed as the solution to the following conditions,

$$\pi = \hat{r}_{t,0}(\hat{h}_t(\pi), x_0),$$
$$1 - \pi = \hat{r}_{0,t}(x_0, \hat{\ell}_t(\pi)).$$

This is done using standard binary search algorithms. In order for this algorithm to work, we assume that $\hat{r}_{0,t}(x_0, \hat{\ell}_t(\pi))$ is decreasing in $\hat{\ell}_t(\pi)$. This assumption is (asymptotically) valid if all goods are normal (i.e. all demand functions are increasing in income).[14]

5.4.2 Adjustments

We make several adjustments to the estimator $\hat{r}_{t,v}(x, y)$ presented above.

Boundary problems The estimator $\hat{r}_{t,v}(x, y)$ has the undesirable property that it may give an estimate strictly between zero and one even when the budget sets $(\mathbf{p}_t, x_t)$ and $(\mathbf{p}_v, x_v)$ do not intersect. In order to see this, assume that $x > \mathbf{p}_t\mathbf{q}$ for all $\mathbf{q}$ for which $\mathbf{p}_v\mathbf{q} = y$, i.e. the budget set defined by price income $(\mathbf{p}_t, x)$ lies strictly above the budget defined by $(\mathbf{p}_v, y)$. In this setting, it may still be the case that the indicator function $\mathbb{1}[x \geq \mathbf{p}_t\mathbf{Q}_{i,v}]$ is zero for some observations i. From this, it follows that the kernel estimator will also be strictly below one although the true value of $\pi_{t,v}(x, y)$ is clearly equal to one.

In order to avoid this boundary problem, we reformulate the probability to be estimated in the following way,

$$\begin{aligned}\int \mathbb{1}[x \geq \mathbf{p}_t\mathbf{q}_v]\, dF(\mathbf{q}_v|x_v = y) &= \int \mathbb{1}[x\mathbf{p}_v\mathbf{s}_v \geq \mathbf{p}_t y\mathbf{s}_v]\, dF(\mathbf{s}_v|x_v = y),\\ &= \int \mathbb{1}[(x\mathbf{p}_v - y\mathbf{p}_t)\,\mathbf{s}_v \geq 0]\, dF(\mathbf{s}_v|x_v = y),\end{aligned}$$

Where $\mathbf{s}_v$ is the random vector of normalised consumption, $\mathbf{s}_v = \mathbf{q}_v/x_v$, and we used

[14] See also Blundell, Browning, and Crawford (2003) for a similar assumption.

the identity $\mathbf{p}_v\mathbf{s}_v = 1$ and the fact that, conditional on $x_v = y$, $\mathbf{q}_v = y\mathbf{s}_v$. If we denote the realisations of $\mathbf{s}_v$ by $\mathbf{s}_{v,i} = \mathbf{q}_{v,i}/x_{v,i}$, we can estimate this using the Nadaraya-Watson estimator,

$$\frac{\frac{1}{nh}\sum_{i=1}^{n}\mathbb{1}\left[(x\mathbf{p}_v - y\mathbf{p}_t)\,\mathbf{s}_{v,i} \geq 0\right] k\left(\frac{\ln(x_{v,i})-\ln(y)}{h}\right)}{\frac{1}{nh}\sum_{i=1}^{n} k\left(\frac{\ln(x_{v,i})-\ln(y)}{h}\right)}.$$

Although this estimator could still over- or underestimate the true proportion slightly, the bias should be considerably less. Also, the estimator has the advantage that it is either zero or one if the two budget lines do not intersect. In order to see this, assume that the two budgets do not intersect, i.e. there is no bundle $\mathbf{q}$ such that

$$\mathbf{p}_t\mathbf{q} = x \text{ and,}$$
$$\mathbf{p}_v\mathbf{q} = y$$

Now, assume that the estimator is somewhere strictly between zero and one. This means that there exist observations i and j such that:

$$(x\mathbf{p}_v - y\mathbf{p}_t)\,\mathbf{s}_{v,i} < 0 \text{ and,}$$
$$(x\mathbf{p}_v - y\mathbf{p}_t)\,\mathbf{s}_{v,j} \geq 0$$

As both left hand sides are continuous functions of the shares vector, we can use the intermediate value function and show the existence of a vector $\mathbf{s}$ such that:

$$(x\mathbf{p}_v - y\mathbf{p}_t)\,\mathbf{s} = 0,$$
$$\Longleftrightarrow \frac{\mathbf{p}_v}{y}\mathbf{s} = \frac{\mathbf{p}_t}{x}\mathbf{s}$$

If we define $\alpha = \frac{y}{\mathbf{p}_v\mathbf{s}} = \frac{x}{\mathbf{p}_t\mathbf{s}}$, and let $\tilde{\mathbf{q}} = \alpha\mathbf{s}$, we have that $\mathbf{p}_t\tilde{\mathbf{q}} = x$ and $\mathbf{p}_v\tilde{\mathbf{q}} = y$, a contradiction.

Semiparametric adjustment We also adjust the kernel estimator $\hat{r}_{t,v}(x,y)$ by including a semi-parametric specification. We have two reasons to do this. First of all, given the data limitations, we would like to allow our estimator to depend on the vector of observed covariates, $\mathbf{a}$, without fully conditioning on each of its values. Next, we need to take into account the fact that total expenditures are probably endogenous. We follow Blundell, Browning, and Crawford (2008), and consider the following semiparametric modification,

$$\mathbb{1}\left[(x\mathbf{p}_t - y\mathbf{p}_v)\,\mathbf{s}_{v,i} \geq 0\right] = g(\ln(x_{v,i}) - \phi(\mathbf{a}'_{v,i}\theta)) + \mathbf{a}'_{v,i}\gamma + \varepsilon_{v,i},$$

where $\mathbf{a}_{v,i}$ be the observed household composition in cross section v for household i. The function $\phi(\mathbf{a}'_{v,i}\theta)$ can be interpreted as the log of a general equivalence scale for the household, and $\mathbf{a}'_{v,i}\gamma$ documents the way in which observable demographic differences across households impact on the left hand side. Similar to Blundell et al. (2008) we use an estimate of the general equivalence scale $\phi(\mathbf{a}'_{v,i}\theta)$ taken from the Organisation for Economic Co-operation and Development (OECD) scales. The semiparametric extension used in this chapter corresponds to the shape invariant specification considered by Härdle and Jerison (1988); Härdle and Marron (1990) and Pinske and Robinson (1995). This specification allows to pool nonparametric estimates of regression curves (e.g. kernel estimates) associated with different households without overly restricting the shape of the curves.

In order to control for the endogeneity of the expenditure level x_v, Blundell, Browning, and Crawford (2008) suggest to use a control function approach based on the two step semiparametric estimator (this is a linearised version of the procedure set out by Newey, Powell, and Vella (1999)). In a first step, we obtain the residuals from a regression of the log of total expenditure on all exogenous variables in the model and on an excluded instrument. We take the log of (equivalent) labour income as an instrument. In the second step, we conduct a semiparametric regression of $\mathbb{1}\left[(x\mathbf{p}_t - y\mathbf{p}_v)\mathbf{s}_{v,i} \geq 0\right]$ on $g(\ln(x_{v,i}) - \phi(\mathbf{a}'_{v,i}\theta))$, $\mathbf{a}'_{v,i}\gamma$ and $\hat{\delta}_{v,i}$, where $\hat{\delta}_{v,i}$ are the residuals from the first stage regression.

5.4.3 Inference on bounds

The methodology outlined in Section 5.3 provides nonparametric bounds on various parameters of interest (e.g. the quantiles of the money metric utility). In the previous section, we have also shown how these bounds can be consistently estimated. However, given that the bounds are based on finite sample estimates, we are confronted with the issue of statistical inference, in particular, the construction of confidence intervals. Given that our estimates only provide bounds, this problem fits in the literature that deals with the construction of confidence intervals for partially identified estimators. We refer to the several recent papers by Imbens and Manski (2004); Chernozhukov, Hong, and Tamer (2007); Stoye (2009); Chernozhukov, Lee, and Rosen (2013) and in particular to the recent paper of Hoderlein and Stoye (2014) who consider the problem of constructing confidence intervals in a setting which is similar to ours.

Given that the main contribution of this study is not on statistical inference, we will only briefly discuss this issue and instead refer to above mentioned papers for more details on how to construct confidence intervals in our setting. As an example of how such construction could look like, consider the case of the quantile $m_t(\pi)$ which gives the πth quantile of the money metric utility $\mu(\mathbf{p}_t, p_0, x_0; \boldsymbol{\nu})$. Using the results above, we know that

$$m_t(\pi) \in \Theta_0 = [\ell_t(\pi); h_t(\pi)]$$

In practice, however, we only have estimates $\hat{\ell}_t(\pi)$ and $\hat{h}_t(\pi)$. There are two kinds of intervals that can be constructed. The first is a confidence interval for the interval Θ_0, i.e. a set CI_α such that,

$$\lim_{nh \to \infty} \Pr\left(\Theta_0 \subseteq CI_\alpha\right) = \alpha.$$

A second kind of interval, CI_α^m, constructs an interval for the quantile $m_t(\pi)$ itself in the sense that,

$$\lim_{nh\to\infty} \Pr\left(m_t(\pi) \in CI^m_\alpha\right) = \alpha.$$

One important result from the literature (see, for example, Imbens and Manski (2004)) is that,

$$\lim_{nh\to\infty} \Pr\left(m_t(\pi) \in CI_\alpha\right) \geq \lim_{nh\to\infty} \Pr\left(m_t(\pi) \in CI^m_\alpha\right) = \alpha$$

As such, any confidence set for the interval, Θ_0, is also a (conservative) confidence interval for the parameter $m_t(\pi)$. An interval CI_α can be constructed in the following way, provided that the estimates of the upper and lower bounds are asymptotically normally distributed with zero asymptotic bias.[15] Let $[\hat{\ell}_t(\pi) - \frac{c_\alpha \hat{\sigma}_l}{\sqrt{nh}}, \hat{\ell}_t(\pi) + \frac{c_\alpha \widehat{\sigma_l}}{\sqrt{nh}}]$ be an asymptotic α% confidence interval for the lower bound $\ell_t(\pi)$ and let $[\hat{h}_t(\pi) - \frac{c_\alpha \widehat{\sigma_h}}{\sqrt{nh}}, \hat{h}_t(\pi) + \frac{c_\alpha \widehat{\sigma_h}}{\sqrt{nh}}]$ be an asymptotic α% confidence interval for the upper bound $h_t(\pi)$, where $\widehat{\sigma_l}$ and $\widehat{\sigma_h}$ are consistent estimates of the standard errors of the asymptotic distribution of the lower and upper bound and where c_α is chosen such that,

$$\Phi(c_\alpha) - \Phi(-c_\alpha) = \alpha,$$

where $\Phi(.)$ is the standard normal probability distribution. Then, using a simple Bonferroni argument, we know that,

$$\left[\hat{\ell}_t(\pi) - \frac{c_\alpha \widehat{\sigma_l}}{\sqrt{nh}}, \hat{h}_t(\pi) + \frac{c_\alpha \widehat{\sigma_h}}{\sqrt{nh}}\right],$$

is a conservative asymptotic α% confidence interval for Θ_0.

Finally, in order to construct confidence intervals for the estimates of the bounds, we notice that the estimates of our bounds are obtained as the maximum or minimum of a number of estimators that are computed using samples from different cross sections. It can be shown that in such cases, the usual bootstrap procedure is not valid (see Andrews (2000) for similar type of examples). In order to obtain asymptotic valid inference we use

[15] For kernel estimators, we could get the bias to converge to zero by undersmoothing.

the subsampling procedure which is presented and discussed in detail by Politis, Romano, and Wolf (1999). Subsampling is similar to bootstrap but the samples taken are smaller and draws are obtained without replacement. The subsampling procedure is valid under very weak assumptions, in particular for extrema estimators such as ours.

5.4.4 Data Description

We illustrate our approach by using a data sample from the Consumer Expenditure Survey (CEX), a repeated cross section. We use data on consumption decisions by US households from 1994 to 2007 (14 years). It is important to note that the consumer expenditures are derived from the diary survey (and not from the interview data). The diary data seem well-suited for (static) demand analysis. First of all, given that we focus on non-durable consumption, which is customary in static demand analysis, information on the purchase of big, durable items is unnecessary. Second, for non-durable commodities, the diary survey invites respondents to indicate their consumption in a two-week period. Because this period is relatively short, respondents should be able to recall their expenditures. We follow Blundell et al. (2008) by focusing our attention to three broad expenditure categories, namely, food, other non-durables and services.[16] As the diary survey reports expenditures on a two-week basis, we convert these to yearly equivalents. Converting two-week expenditures to yearly data poses an important problem of seasonality. Therefore, we deseasonalise using a dummy regression approach. Specifically, the expenditures on each category (reported for two weeks) are regressed on month dummies. Residuals from this regression (which can be interpreted as the variation in expenditures which can not be explained by seasonality or by months) are added to the mean expenditures for each category in order to construct deseasonalised expenditures. Observations with negative total expenditures are dropped. As mentioned above, we also take into account that variation in expenditures can be driven by the household composition, e.g. the number of adults or the number of kids living in

[16]See Appendix 5.A for a list of the different goods used for the construction of the aggregates.

the family. Therefore, we deflate total expenditures as well as total income by an OECD equivalence scale.

For the empirical analysis, we restrict attention to (i) households who have completed the two-week diary, (ii) households who are not living in student housing, (iii) households who are vehicle owners (to include fuel expenses), (iv) households where both members work at least 17 hours, (v) households in which both members are not self-employed, (vi) households in which the age of the reference person is at least 21 and finally we restrict attention to (vii) households that consist of a husband, a wife and possibly children. As a final step we also remove some outlier observations.[17] On average, we are left with 2163 observations per cross-section with a minimum of 1775 observations in 1994 and a maximum of 2379 observations in 2007. The upper panel in Figure 5.3 plots the evolution of the mean consumption shares of the three goods over the considered periods. The lower panel plots the evolution of prices obtained from the Bureau of Labor Statistics. Finally, Figure 5.4 sets out the evolution of different percentiles of the distribution of total expenditures across all years.

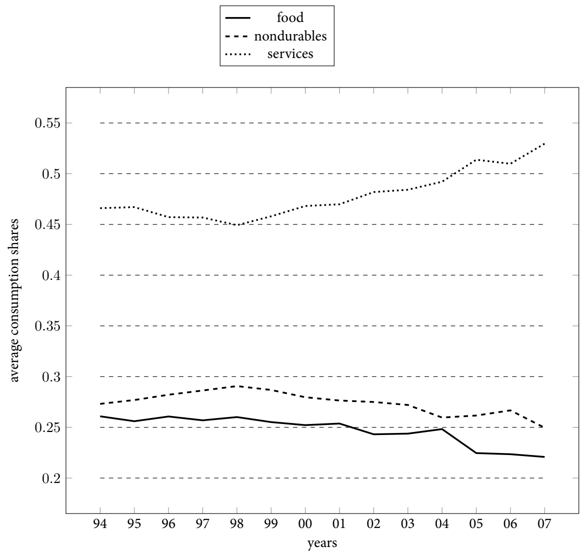
5.4.5 Empirical results

In this section, we provide the results of several exercises. Due to limited space, we need to restrict our analysis to some particular base years and some reference income levels. Additional results are available from the authors upon request.

Bounds on the mean cost of living Let us first show how our bounds perform with respect to the computation of the mean of the Laspeyres-Konüs cost of living index,

$$\int \frac{\mu(\mathbf{p}_t, \mathbf{p}_0, x_0; \boldsymbol{\nu})}{x_0} dF(\boldsymbol{\nu}) = \frac{1}{x_0} \int \mu(\mathbf{p}_t, \mathbf{p}_0, x_0; \boldsymbol{\nu}) dF(\boldsymbol{\nu}).$$

[17] In particular, we removed observations for which rescaled total expenditures or expenditure shares are not within 3 standard deviations from the mean and observations for which rescaled total expenditures are among the 5 per cent lowest or 5 per cent highest expenditures or for which the expenditure shares on are close to 0.



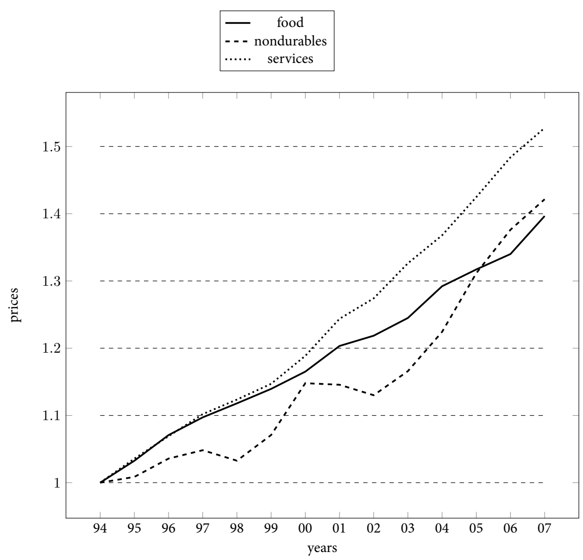


Figure 5.3: Evolution of average consumption shares and prices

The Laspeyres-Konüs price index measures the income that one would need, relative to the income in period 0, in order to be equally well off as in the initial period. We take 1994 as the reference year which means that $\mathbf{p}_0$ corresponds to the price vector in the year 1994. We

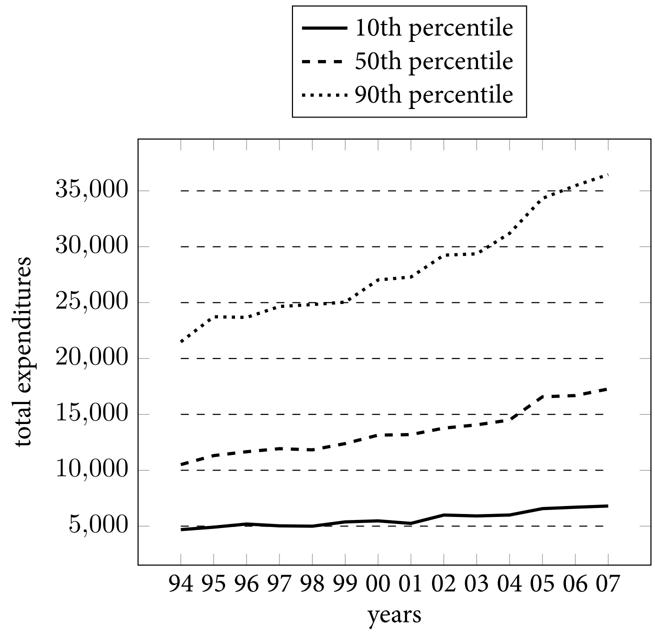


Figure 5.4: Evolution of 10th, 50th and 90th percentiles of total expenditures

choose x_0 as the (OECD equivalence scale deflated) median expenditure level in 1994. The bounds on the cost of living that we obtain using our procedure are given in the last column of Table 5.1. The table also reports values for various other prices indices like the Laspeyres (L), the Paasche (P) and the Tornqvist price index (T).[18] We also provide information on three other nonparametric bounds. The first are the Lerner bounds which are obtained from the fact that:

$$\min_j \left\{ \frac{p_{t,j}}{p_{0,j}} \right\} \leq \frac{\mu(\mathbf{p}_t, \mathbf{p}_0, x_0)}{x_0} \leq \max_j \left\{ \frac{p_{t,j}}{p_{0,j}} \right\}.$$

The bounds by Pollak (1971) improve upon this by replacing the upper bound by the Laspeyres price index.

$$\min_j \left\{ \frac{p_{t,j}}{p_{0,j}} \right\} \leq \frac{\mu(\mathbf{p}_t, \mathbf{p}_0, x_0)}{x_0} \leq \frac{\mathbf{p}_t \mathbf{q}_0}{x_0}.$$

[18] These are computed on the basis of nonparametric Engel curve estimates.

The second to last column gives the bounds that are obtained by using the procedure set out by Blundell, Browning, and Crawford (2003). This method first estimates nonparametric Engel curves and subsequently uses these estimates in combination with revealed preference restrictions to establish nonparametric bounds. We would like to emphasise that there is a clear conceptual difference between the bounds of Blundell, Browning, and Crawford (2003) (and Pollak), and ours. Their procedure provides bounds on the cost of living that correspond to some kind of 'representative individual' whose demand functions equal the mean demand functions over the population. Our bounds, on the other hand, correspond to bounds on the mean cost of living over all households within the population. Although it is reassuring to see that both procedures give very similar results, this does not have to be the case in general.

	Price indices			Nonparametric Bounds			
year	L	P	T	Lerner	Pollak	BBC	bounds
1994	1.0000	1.0000	1.0000	[1.0000, 1.0000]	[1.0000, 1.0000]	[1.0000, 1.0000]	[1.0000 1.0000]
1995	1.0275	1.0271	1.0273	[1.0086, 1.0357]	[1.0086, 1.0275]	[1.0250, 1.0275]	[1.0249 1.0292]
1996	1.0604	1.0596	1.0600	[1.0358, 1.0708]	[1.0358, 1.0604]	[1.0591, 1.0604]	[1.0574 1.0621]
1997	1.0860	1.0844	1.0852	[1.0483, 1.1019]	[1.0483, 1.0860]	[1.0830, 1.0860]	[1.0819 1.0875]
1998	1.0972	1.0929	1.0951	[1.0327, 1.1236]	[1.0327, 1.0972]	[1.0932, 1.0972]	[1.0900 1.0983]
1999	1.1242	1.1212	1.1227	[1.0709, 1.1470]	[1.0709, 1.1242]	[1.1205, 1.1242]	[1.1180 1.1256]
2000	1.1716	1.1712	1.1714	[1.1480, 1.1886]	[1.1480, 1.1716]	[1.1689, 1.1716]	[1.1692 1.1739]
2001	1.2066	1.2048	1.2057	[1.1456, 1.2437]	[1.1456, 1.2066]	[1.2025, 1.2066]	[1.2025 1.2086]
2002	1.2206	1.2154	1.2181	[1.1301, 1.2742]	[1.1301, 1.2206]	[1.2143, 1.2201]	[1.2122 1.2222]
2003	1.2618	1.2562	1.2591	[1.1659, 1.3263]	[1.1659, 1.2618]	[1.2556, 1.2607]	[1.2534 1.2636]
2004	1.3094	1.3066	1.3080	[1.2243, 1.3679]	[1.2243, 1.3094]	[1.3048, 1.3089]	[1.3042 1.3115]
2005	1.3666	1.3698	1.3682	[1.3115, 1.4247]	[1.3115, 1.3666]	[1.3648, 1.3665]	[1.3658 1.3706]
2006	1.4181	1.4206	1.4194	[1.3400, 1.4839]	[1.3400, 1.4184]	[1.4157, 1.4178]	[1.4164 1.4202]
2007	1.4655	1.4679	1.4667	[1.3966, 1.5276]	[1.3966, 1.4655]	[1.4633, 1.4655]	[1.4644 1.4691]

Table 5.1: Bounds on the mean Laspeyres Konüs cost of living index

Distribution of the cost-of-living Let us now have a look at the bounds on the quantiles of this cost of living index over the population. Figure 5.5 provides bounds on the quantiles of the Laspeyres-Konüs cost of living index, for the 10th (red), 50th (black) and 90th (blue) percentile. Upper and lower bounds on a particular quantile are presented by the same color.

Again the base year is 1994 and the reference income is given by the median expenditure level in this year. In general, we see that the bounds on the quantiles are quite narrow. The width of the distribution for a particular year depends to a large extent on the difference in relative slopes between the base year ($\mathbf{p}_0/x_0$) and the evaluation year ($\mathbf{p}_t/x_t$). The closer the relative prices, the narrower the difference between the largest and smallest cost of living for the particular year. The reason is that the distribution is naturally bounded between the minimum and maximum values of y/x_0 for which the budget hyperplanes corresponding to $(\mathbf{p}_t, y)$ and $(\mathbf{p}_0, x_0)$ do not intersect. We see that the distribution is narrow in the year 2000 and the widest in the year 2002 giving differences in cost of living up to more than 5 percentage points between the 10th and 90th percentile. One noticeable feature about the figure is that there seems to be a considerable amount of heterogeneity in the population although the width of the distribution remains more or less constant for the latter 5 years. Finally, we present confidence intervals (from subsampling) for the results in Figure 5.5 (and for the 30th and 70th percentile) in Appendix 5.B.

Figure 5.6 gives another illustration of the kind of questions that can be answered given the framework in this paper. The figure gives bounds on the average of the Laspeyres-Konüs cost of living for different starting quantiles,

$$\int \frac{\mu(\mathbf{p}_t, \mathbf{p}_0, x_{0,i}, \boldsymbol{\nu})}{x_{0,i}} dF(\boldsymbol{\nu}).$$

Here, $x_{0,i}$ represents the income at the ith quantile of the income distribution in 1994, $\mathbf{p}_0$ are the prices in 1994 and $\mathbf{p}_t$ is the price vector for 2007. The figure gives an idea of the average price increase (over the heterogeneous population) for households starting at different quantiles of the income distribution in 1994. On average one sees an increase over the quantiles, which means that (on average) the cost of living for households starting at the lower end of the income distribution in 1994 was lower than for household starting at the higher end of the income distribution. In other words, the households that started at the

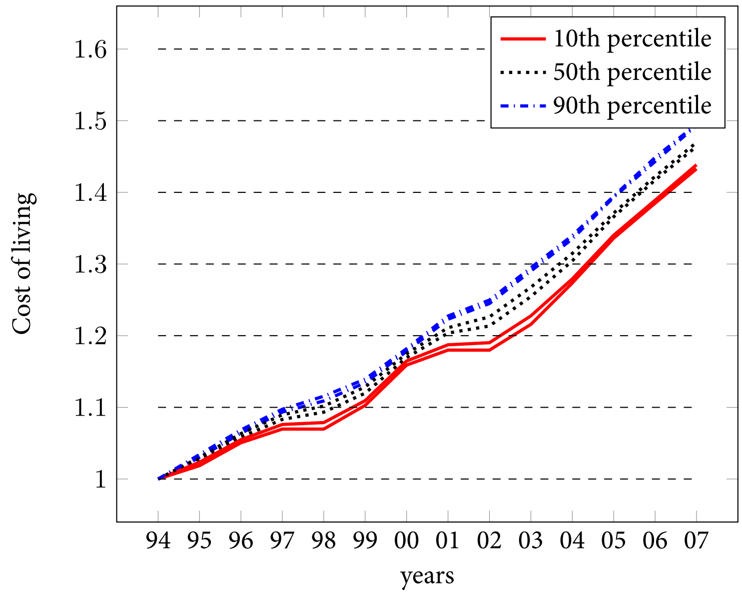


Figure 5.5: Distribution of the cost of living

lower end of the income distribution had (on average) a lower increase in the cost of living. Also, notice that the upper bound for the lowest quantile is below the lower bound for the upper quantile. This shows that the average cost of living values are significantly different (although the numbers are very close to each other in absolute terms).

Distribution of the compensating variation Figure 5.7 shows the distribution of the compensating variation,

$$x_t - \mu(\mathbf{p}_t, \mathbf{p}_0, x_0; \boldsymbol{\nu})$$

Here, x_0 is taken to be the median income in 2000 and x_t is the median income in cross section t. This compensating variation gives the difference between the median income in year t and the minimum income that would be necessary in order to obtain the welfare level at budget $(\mathbf{p}_0, x_0)$. Values above zero indicate a welfare gain for a household at the median income in year t compared to a household at the median income in year 1994. We see that

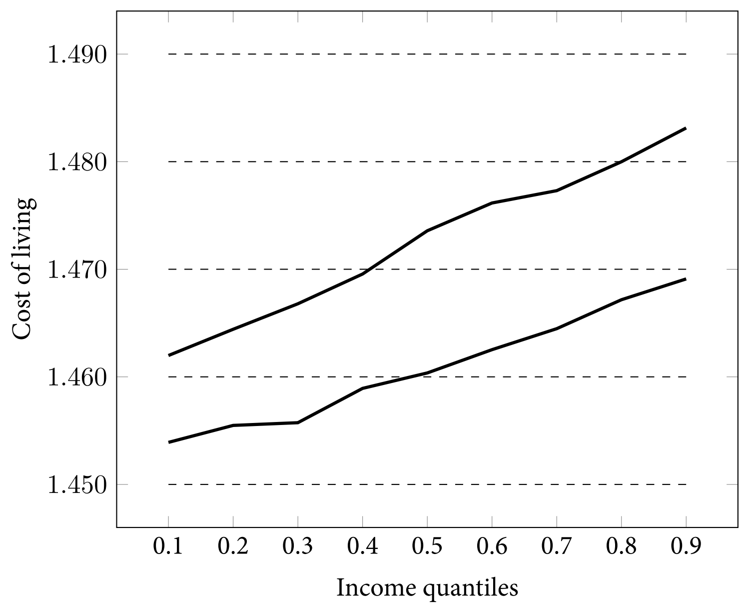


Figure 5.6: Change in mean cost of living 1994-2007 for different starting quantiles of income

all quantiles are below zero for the years 1994-1999 and 2001 and quantiles are above zero for the years 2005-2007. This corresponds to the significant increase in total expenditures after 2004, presented in Figure 5.4.

Once again, there seems to be quite a lot of heterogeneity present in the population. For many years, the range between the 10th and 90th percentile is around \$400 per year which is substantial.

Distribution of demand As a last exercise, let us have a look at the bounds on the demanded consumption shares for counterfactual price regimes. To keep focus, we restrict ourselves to the computation of bounds for the own price effect for the food aggregate. We construct normalised prices by dividing all cross sectional prices by the median income in the corresponding year, and we take the mean of these normalised prices as a reference point. The reference income level x_0 is set at 1. We let the price for food range from 0.95

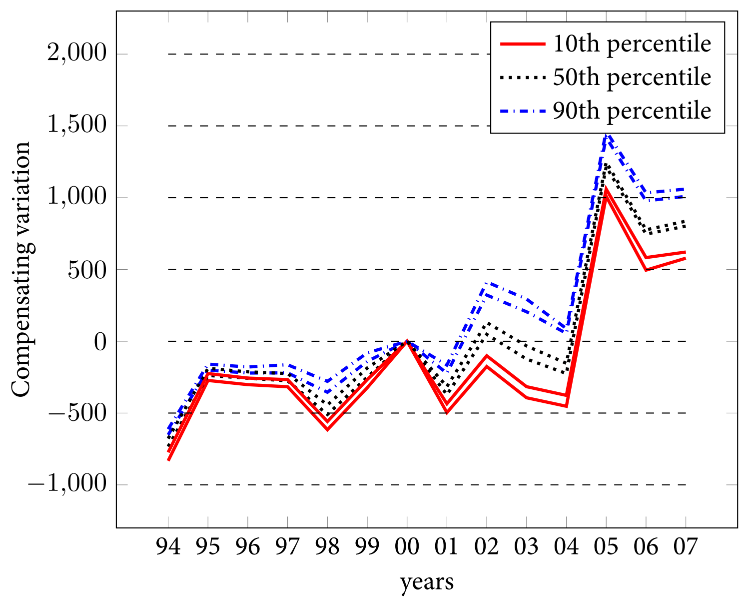


Figure 5.7: Distribution of compensating variation, baseyear 2000

times its reference level to 1.05 times its reference level. The prices of all others goods are held constant. Figure 5.8 presents the results for three quantiles, the 10th (in red), the median (in black) and the 90th (in blue). Again we see a lot of heterogeneity in the demand curves over the population although the price responses look very similar across the three quantiles. Interestingly, the bounds seem to allow for Cobb Douglas preferences. It is possible to construct a curve of expenditure shares (in function of prices) which is relatively flat, indicating that expenditure shares are independent of prices.

As is customary in revealed preference analysis, it is only possible to construct bounds on the counterfactual demands for prices in the convex support of the observed prices. This explains the large and simultaneous drop of all lowerbounds at the 3% price increase.

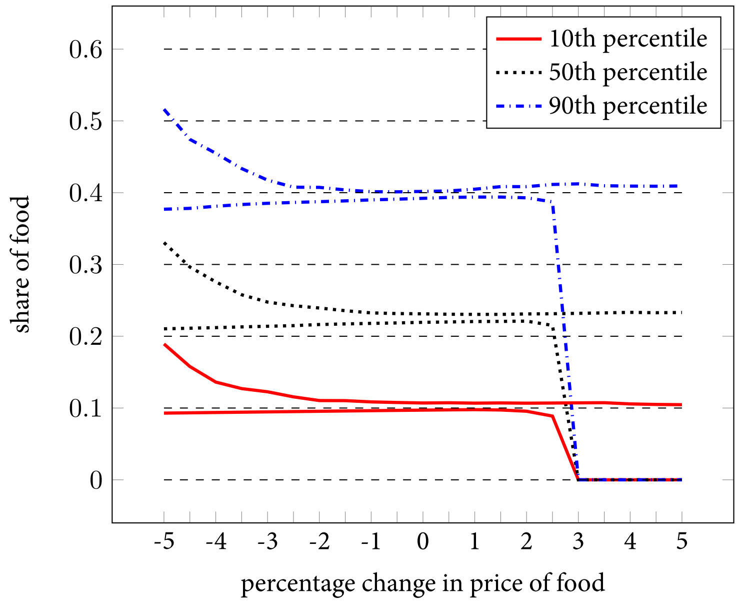


Figure 5.8: Bounds on the demand shares

5.5 Conclusion

In this paper, we used elementary revealed preference techniques together with nonparametric estimation techniques in order to bound the distribution of the money metric utility and the demand functions over a population of heterogeneous households. Our methodology has two attractive features. First of all, the results are entirely nonparametric which means that they are not dependent on any functional form imposed on the underlying utility functions. Second, we impose minimal conditions on the structure of the individual, unobserved heterogeneity. When we apply our techniques to data from the US consumer expenditure survey, we find that our method generates narrow and informative bounds on the quantiles of the money metric utility function. Our results also demonstrate that individual heterogeneity creates considerable variation in welfare between households in the population (conditional on the same level of expenditure). We also demonstrate how our results can be used to obtain informative bounds on the distribution of the demand functions in counterfactual price-income situations.

There are several avenues for follow up research. First of all, we only briefly touched upon the highly relevant topic of statistical inference. However, given that our data is obtained from a random sample, measurement error and small sample biases influence our bounds, and statistical inference becomes relevant. Next, it would be interesting to see how our methodology extends to discrete choice settings. One way to incorporate discrete choices would be to consider a setting where individuals make discrete choices in addition to continuous choices. Many of the results from this paper readily extend to such setting. Alternatively one could imagine a setting where all choices are discrete (see Manski (2007) and Sher, il Kim, Fox, and Bajari (2011) for a theoretical account of stochastic revealed preferences recovery in such setting). It would be interesting to look how the methodology developed in this paper transfers to such discrete choice setting. Finally, it would be interesting to see how other (more strict) stochastic revealed preference axioms that explicitly take into account transitivity may even further improve our bounds.

5.A Construction of aggregates

Food is an aggregate of cereals, bakery products, beef, pork, poultry, seafood, other meat, eggs, milk products, other dairy products, fresh fruit, fresh vegetables, processed fruit, processed vegetables, sweets, fat and oils, non-alcoholic beverages, prepared food, snacks and condiments.

Other non-durables contain expenditures on alcohol consumption, tobacco, clothes (for all household members), footwear, reading material, stationery, school supplies, cleaning products, garden supplies, household textile, non-durable housewares, medical products, personal care products, audio-visual equipment, recreational goods, pet goods and vehicle expenses.

Services include utilities, media bills, repair services, insurance, postal services, gasoline, vehicle expenses (services), public transportation, medical care services, personal care services, recreational services, home services, rental services, membership fees, school fees, other fees, pet services and care services.

5.B Confidence intervals

To compute the Bonferroni intervals we followed a subsampling procedure (in line with Politis, Romano, and Wolf (1999)). Subsampling is similar to the bootstrap procedure but the samples are smaller and the draws are obtained without replacement. Consider a dataset of size n and an estimator $\hat{g}_n$ which converges at a rate such that $\tau_n(\hat{g} - g)$ converges to a non-degenerate asymptotic distribution for $n \to \infty$. In our case, $\tau_n = \sqrt{nh_n}$ where h_n is the bandwidth. The subsampling procedure proceeds by taking (without replacement) subsamples of size m and computing the associated estimator g^*_m. Then, under very weak conditions, it can be shown that for $m \to \infty$, $m/n \to 0$ and $\tau_m/\tau_n \to 0$ as $n \to \infty$, the statistic $\tau_m(g^* - \hat{g})$ converges to the same asymptotic distribution as $\tau_n(\hat{g} - g)$.

We apply the subsampling approach to the distribution of the (Laspeyres-Konüs) cost of living index using 999 subsamples of size $m \approx \sqrt{n}$. Table 5.2 presents our original estimates of upper and lower bounds on the quantiles of the Laspeyres-Konüs cost of living index and the corresponding asymptotic 95% confidence intervals for the setting in Figure 5.5.

year	10th percentile		30th percentile		50th percentile		70th percentile		90th percentile	
bounds 1994	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
CI 1994	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
bounds 1995	1.019	1.023	1.023	1.028	1.025	1.030	1.028	1.032	1.030	1.034
CI 1995	1.018	1.025	1.022	1.029	1.025	1.031	1.027	1.033	1.030	1.035
bounds 1996	1.051	1.055	1.055	1.061	1.058	1.063	1.061	1.066	1.064	1.069
CI 1996	1.049	1.056	1.054	1.061	1.058	1.064	1.060	1.066	1.063	1.069
bounds 1997	1.070	1.076	1.078	1.085	1.083	1.090	1.088	1.093	1.093	1.097
CI 1997	1.068	1.078	1.077	1.086	1.082	1.091	1.087	1.094	1.092	1.098
bounds 1998	1.070	1.079	1.085	1.094	1.093	1.102	1.100	1.108	1.109	1.116
CI 1998	1.066	1.082	1.081	1.096	1.091	1.103	1.098	1.110	1.107	1.117
bounds 1999	1.103	1.109	1.113	1.122	1.120	1.128	1.127	1.134	1.134	1.139
CI 1999	1.100	1.112	1.111	1.124	1.118	1.130	1.126	1.135	1.133	1.141
bounds 2000	1.159	1.165	1.165	1.170	1.169	1.175	1.173	1.178	1.178	1.182
CI 2000	1.158	1.166	1.164	1.171	1.168	1.176	1.172	1.179	1.177	1.183
bounds 2001	1.180	1.187	1.193	1.202	1.204	1.211	1.213	1.219	1.223	1.227
CI 2001	1.177	1.190	1.191	1.204	1.201	1.213	1.210	1.220	1.221	1.229
bounds 2002	1.180	1.190	1.199	1.213	1.214	1.226	1.226	1.237	1.245	1.250
CI 2002	1.177	1.195	1.195	1.216	1.211	1.229	1.223	1.239	1.241	1.252
bounds 2003	1.216	1.228	1.240	1.252	1.254	1.268	1.268	1.280	1.290	1.295
CI 2003	1.212	1.232	1.237	1.255	1.252	1.270	1.265	1.282	1.286	1.297
bounds 2004	1.273	1.279	1.292	1.300	1.304	1.315	1.317	1.326	1.335	1.340
CI 2004	1.270	1.283	1.289	1.304	1.302	1.317	1.315	1.328	1.332	1.343
bounds 2005	1.336	1.340	1.354	1.358	1.366	1.371	1.378	1.382	1.393	1.395
CI 2005	1.333	1.343	1.351	1.361	1.363	1.373	1.375	1.384	1.390	1.397
bounds 2006	1.385	1.389	1.403	1.408	1.417	1.422	1.431	1.435	1.443	1.449
CI 2006	1.381	1.393	1.400	1.411	1.414	1.425	1.427	1.437	1.440	1.452
bounds 2007	1.433	1.439	1.452	1.457	1.464	1.470	1.475	1.481	1.493	1.495
CI 2007	1.430	1.442	1.449	1.460	1.461	1.472	1.473	1.483	1.490	1.498

Table 5.2: Confidence bounds on the distribution of the Laspeyres-Konüs cost of living index - bounds: sample estimates of lower and upper bounds on cost of living, CI: confidence interval with lower and upper bounds on cost of living

Part V

General Conclusion

Let me briefly review the contributions set out in this dissertation, and subsequently propose avenues for follow-up research.

I have presented several extensions of the revealed preference method. The extensions clearly show that the revealed preference methodology cannot be confined to one particular theory or one particular application. On the theoretical level, I have shown how Afriat's Theorem no longer holds in a finite choice set-setting. I also extended the revealed preference approach to allow for non-classical (value-dependent) preferences and I considered a revealed preference approach to study externalities in consumption. On a more applied level, I have shown how to use revealed preference in an experimental setting with finite choice sets. I have also explained how revealed preference can deal with unobserved heterogeneity when panel data are unavailable. All of my contributions build on the so called 'restricted domain version' of revealed preference, developed by Afriat (1967), Diewert (1973) and Varian (1982). These authors showed how revealed preference tests can provide information on the consumer's degree of rationality, his or her preferences and even his or her welfare without restricting demands or utility functions in any sense, and when only a limited number of price-demand observations are available. I have positioned my contributions along the lines of three main themes in demand analysis (in general) and revealed preference (in particular): *testing*, *identification* and *prediction*.

First of all, the revealed preference approach can be used to *test* consistency with the hypothesis of utility maximisation subject to a linear budget. Testing the utility maximisation hypothesis is clearly non-trivial, because the rationality assumption lies at the basis of many identification and estimation strategies. Moreover, rationality is relevant in itself given that irrational behaviour is wasteful behaviour. Standard revealed preference axioms, such as the Generalised Axiom of Revealed Preference (GARP), test rationality while assuming that consumers make choices from linear budget sets. Nonetheless, there are many settings in which choice sets are inherently finite. First, many real-life consumption decisions involve integer quantities, or distinct alternatives with different characteristics. Second, experiments involving children are typically based on finite choice sets. After all, letting respondents choose from finite choice sets allows the researcher to restrict attention to the pure concept of rationality, whereas letting respondents choose from linear budget sets also involves practical problems of budget exhaustion and tedious calculations which are, in particular, complex tasks for children.

In Chapter 1, I have used new experimental data on children's consumption decisions from finite choice sets. The design of our experiment is similar to the experiment by Harbaugh, Krause and Berry (2001) but our experiment involves choices between 3 commodities instead of 2. This more complicated design, in combination with the finite nature of the choice set, also implied that the GARP is no longer a necessary and sufficient test for rationality. I therefore discussed an alternative revealed preference approach to analyse choices from finite choice sets. However, the focus of Chapter 1 was on testing rationality and explaining the drivers of rationality among children. Towards this end, I used information on various child characteristics. The results mainly suggested that younger children are less likely to be rational. I have also found that it is important to take the multidimensional nature of intelligence into account when explaining rationality.

In Chapter 2, I have further investigated the implications of a finite choice-set setting on the equivalence between different rationalisability concepts. Specifically, rationalisability by a weakly monotone utility function, rationalisability by a strongly monotone utility function, rationalisability by a weakly monotone and concave utility function and rationalisability by a strongly monotone and concave utility function are no longer equivalent when choice sets are finite. I have shown that the rationality results obtained by Harbaugh, Krause and Berry (2001) and the ones obtained in Chapter 1 strongly depend on the characterisation at hand. I also investigated assumptions on the choice sets under which the characterisations are still equivalent. These assumptions explain why the GARP could be applied by Harbaugh, Krause and Berry (2001) whereas alternative axioms were required to analyse the choices in Chapter 1. At a more general level, the results from Chapter 2 built a bridge between the 'extended domain version' of revealed preference in the spirit of Arrow (1959) and the 'restricted domain version' of revealed preference following Afriat (1967), Diewert (1973) and Varian (1982). Indeed, I studied consumption decisions from finite choice sets but I used these choices in order to test rationalisability by standard (monotone and/or concave) utility functions.

Second, the revealed preference method can *identify* elements of the preference structure of economic agents. For neo-classical preferences, recovery typically focuses on the 'revealed better' or 'revealed worse' regions in the indifference maps of individuals. In Part 2, I have studied identification for alternative types of preferences. In particular, I allowed for externalities in consumption (preferences for others' consumption, Chapter 3) and diamond effects (preferences for value, Chapter 4).

In Chapter 3, I have proposed 'selfishness' parameters to capture individuals' willingness to pay for own consumption (vis-a-vis the other's consumption). I embedded these parameters in the revealed preference characterisation of the collective model of consumer behaviour, which assumes that collective decisions lead to a Pareto efficient outcome. Hence, instead of relaxing the assumption of Pareto efficiency (see e.g. the revealed preference characterisation of the non-cooperative model by Cherchye et al. (2011b)), the 'selfishness' parameter relaxes the assumption of purely self-interested consumers. After all, there is much evidence that individuals not only care about the own consumption. I have applied the method to experimental data on joint consumption decisions by children. Information on the intra-dyad allocation of private goods allowed to bound the selfishness parameter. I found that selfishness is related to various dyad characteristics, such as the level of friendship between group members. As expected, children had stronger preferences for the other's consumption (i.e. a lower selfishness parameter) when the group partner was a friend. On a more general level, my approach implemented an idea put forward by Rabin (2013): the selfishness parameter determines whether the collective model reduces to the egoistic model (according to which consumers care only about own consumption) or whether externalities are allowed. As such, it easily defined portable extensions of an existing model.

In Chapter 4, I extended the GARP to analyse individual consumers who have preferences for the value of a purchase. This corresponds to the so called diamond effect, introduced by Ng, 1987. The purpose of this chapter was to bring the conjecture of value-(and price-)dependent preferences to the data, without making prior assumptions on the form of utility functions. Moreover, by following a revealed preference approach, I could analyse diamond (and conspicuous consumption) effects without having to assume that different people have homogeneous preferences for value. In particular, I have modified the GARP to test for utility functions with both quantities and value as arguments. The newly proposed test allowed to distinguish between preferences for material consumption and preferences for the value associated with consumption (i.e. the diamondness). The results suggested that there are strong preferences for value and that these preferences are driven by conspicuous consumption effects. Specifically, commodities with higher visibility scores were associated with higher diamondness values. This insight might motivate differentiated tax rates for commodities like alcohol, luxury clothing, jewellery, etc (Ng, 1987). Again, this contribution fits in Rabin's (2013) PEEM approach, by adding psychological realism to a model while still retaining testable implications.

Finally, Chapter 5 differed from the other chapters in two ways. First, I focused on the estimation and prediction of welfare measures and demand correspondences rather than on the consistency or identification of preferences. Second, I have applied the method to cross-sectional data (from the US Consumer EXpenditure survey) rather than panel data. This leads to issues of unobserved heterogeneity because it is no longer possible to analyse each agent separately. The main contribution of Chapter 5 is the application of the stochastic revealed preference conditions (initiated by McFadden and Richter (1971) and Falmagne

(1978)) to derive bounds on the distribution of welfare and the distribution of demand. Insight into the distribution of welfare and demand across the population is important from a policy perspective. It allows policy makers and researchers to study the gap between households with the highest level of welfare and households who are worst off. One can reasonably argue that extreme variation in welfare across households is unfavourable. The method also allows policy makers to estimate and/or predict the change in welfare that results from an income or price change. The method presented in Chapter 5 imposes no structure on the utility functions of consumers and treats unobserved heterogeneity in the most general (non-additive) way.

In this dissertation, I have discussed the versatility of the revealed preference method. The scope for future research is obviously very large. In the final paragraphs, I will restrict attention to extensions that could follow from the contributions presented in this thesis.

First of all, I have treated children as decision makers in Chapters 1, 2 and 3. Children could choose from different sets of alternatives. The results clearly suggested that young children are unable to consume rationally. The negative relationship between children's ability to make rational decisions on the one hand and their age on the other hand was prominent. For this reason, parents typically make decisions on behalf of the (young) children. Child well-being can then be treated as an intra-household (and domestically produced) public good. Blundell et al. (2005) and Cherchye et al. (2012), for instance, assume a domestic production function which maps parental time and resources allocated to children on a variable that indicates the child's utility. Both studies use a parametric form for this production function. An interesting avenue for future research is therefore the development of a fully nonparametric counterpart for these models. Such study would require data on the time use of parents and the intra-household allocation of consumption goods. This type of data is available, for instance, in the LISS (Longitudinal Internet Studies for the Social sciences) panel.

Second, I have presented a general method to recover willingness to pay for others' consumption in Chapter 3. The proposed 'selfishness' parameter allowed us to compare the egoistic model of collective consumption with more general collective models that incorporated strong externalities in consumption. I have not distinguished between different sources of externalities. There are various reasons why individuals could care about the consumption of others: altruism, preferences for equality, preferences for the material pay-off of the least well off, etc. The focus of Chapter 3 was on the measurement of other-regarding preferences, in general. Future research could impose additional restrictions on the utility functions which, in combination with a specifically tailored experiment, would shed light on the precise motivation underlying these other-regarding preferences. Another extension could focus on the implicit assumption of Pareto efficient decision-making. The collective model used in Chapter 3 assumes that the joint decisions lead to a Pareto efficient outcome, which implies that the sum of marginal willingness' to pay for one individual's 'assignable' good is equal to the market price in equilibrium. In my opinion, it would also be interesting to investigate what happens if the Pareto efficiency requirement is replaced with the assumption that the outcome corresponds to a Nash equilibrium. Towards this end, one could apply the so called noncooperative model. According to this model, group members are not able or willing to coordinate to achieve the 'optimal' level of consumption, for instance because they conceal their true marginal willingness to pay for the good. Interestingly, Cherchye et al. (2011b) provided a revealed preference characterisation of noncooperative group consumption. This could be a starting point to incorporate preferences for others' consumption and externalities in a setting where the assumption of Pareto efficiency does not hold.

A third avenue for future research is to further investigate the link between the visibility of commodities and the corresponding conspicuous consumption effects. In Chapter 4, I found more outspoken preferences for value associated with the more visible commodity groups. On the basis of these results, one could expect that particular consumers signal their wealth by purchasing visible (and expensive) goods. In this respect, a model which combines the rationality assumption with a signalling equilibrium might significantly contribute to the understanding of individuals' decisions. One could, for instance, define individual utility functions over material consumption on the one hand and a perceived status outcome on the other hand. Each consumer would then form beliefs regarding the status outcome of another consumer based on the position of the others' level of visible consumption in the distribution of visible consumption across the population. Beside rationality, a signalling equilibrium would also require that the consumers' beliefs are (Bayesian) consistent. The combination of a signalling model and the revealed preference methodology could construct a powerful framework to study conspicuous consumption effects.

The fourth avenue for further research is to combine the insights from Chapter 2 with the contributions in Chapter 5. Indeed, one could investigate the implications of stochastic revealed preference when choice sets are finite. I have already argued that finite choice sets occur naturally in many settings. An interesting paper in this respect is by Bhattacharya (2014). The author presents a finite-choice analog to the paper of Hausman and Newey (2013) by showing that various welfare measures are nonparametrically point-identified (rather than set-identified) for binary and multinomial choice.

Finally, in my General Introduction I have explained that revealed preference is a nonparametric methodology to study demand. Similar to other econometric methods, the revealed preference approach is typically concerned with Type I and Type II error. On the one hand, the revealed preference method should not reject the (null) hypothesis of utility maximisation whenever it holds true. On the other hand, irrational consumption data - which are for example simulated by letting a computer draw random bundles - should not pass the corresponding revealed preference test. The magnitude of Type II error in revealed preference is estimated by so called discriminatory power measures, i.e. the higher discriminatory power, the less likely that random data sets will pass the conditions of a model. Together with pass rates, discriminatory power provides information on the empirical performance of a revealed preference model. However, the assessment procedure itself is not independent of methodological choices.

First, there are many ways to simulate random bundles. In this dissertation, I have simulated random data sets both by drawing budget shares from a uniform distribution (Bronars' approach, 1987) and by randomly combining the observed budget shares of different, heterogeneous consumers (bootstrap approach). Andreoni, Gillen, and Harbaugh (2011) present an insightful overview of different methods to estimate discriminatory power.

Second, throughout the dissertation I summarised pass rates and power by using a measure of predictive success. The predictive success provides a direct comparison of the pass rate of observed data and the pass rate of random, simulated data (which is one minus power). However, it is clear that predictive success information is no substitute for the separate reporting of pass rates and power. Moreover, the 'additive' specification of the predictive success measure is only one possible way to summarise pass rates and power. In this respect, I refer to Selten (1991) and Beatty and Crawford (2011) who have shown that the additive predictive success measure satisfies some desirable properties. One of these properties is aggregability. Aggregability implies additivity of individual predictive success measures across the sample. According to Selten's Theorem, all other measures that satisfy these properties are simple linear transformations of the additive predictive success measure.

The two arguments discussed above illustrate that the assessment procedure itself is based on various implicit assumptions. It is clear that the further development of the revealed preference approach as a successful methodology depends on the development of convincing measures of empirical performance.

Bibliography

Afriat, S. N., 1967. The construction of utility functions from expenditure data. International Economic Review 8, 67–77.

Afriat, S. N., 1972. Efficiency estimation of production functions. International Economic Review 13, 568–598.

Afriat, S. N., 1973. On a system of inequalities in demand analysis: An extension of the classical method. International Economic Review 14, 460–472.

Andreoni, J., Gillen, B. J., Harbaugh, W. T., 2011. The power of revealed preference tests: Ex-post evaluation of experimental design. Tech. rep.

Andreoni, J., Miller, J., 2002. Giving according to garp: An experimental test of the consistency of preferences for altruism. Econometrica 70, 737–753.

Andrews, D. W. K., 2000. Inconsistency of the bootstrap when a parameter is on the boundary of the parameter space. Econometrica 68, 399–405.

Apps, P. F., Rees, R., 1988. Taxation and the household. Journal of Public Economics 35, 355–369.

Arrow, K. J., 1959. Rational choice functions and orderings. Economica 26, 121–127.

Bandyopadhyay, T., Dasgupta, I., Pattanaik, P. K., 1999. Stochastic revealed preference and the theory of demand. Journal of Economic Theory 84, 95–110.

Bandyopadhyay, T., Dasgupta, I., Pattanaik, P. K., 2002. Demand aggregation and the weak axiom of stochastic revealed preference. Journal of Economic Theory 107, 483–489.

Bandyopadhyay, T., Dasgupta, I., Pattanaik, P. K., 2004. A general revealed preference theorem for stochastic demand behavior. Economic Theory 23, 589–599.

Banks, J., Blundell, R., Lewbel, A., 1997. Quadratic engel curves and consumer demand. The Review of Economics and Statistics 79, 527–539.

Barberá, S., Pattanaik, P. K., 1986. Falmagne and the rationalizability of stochastic choices in terms of random orderings. Econometrica 54, 707–715.

Basmann, R., Molina, D., Slottje, D., 1988. A note on measuring veblen's theory of conspicuous consumption. Review of Economics and Statistics 70, 531–535.

Basu, K., 1980. Revealed Preference of Government. Books on Demand.

Beatty, T. K. M., Crawford, I. A., 2011. How demanding is the revealed preference approach to demand. American Economic Review 101, 2782–2795.

Becker, G. S., 1962. Irrational behavior and economic theory. Journal of Political Economy 70, 1–13.

Becker, G. S., 1974. A theory of social interactions. Journal of Political Economy 82, 1063–1093.

Becker, G. S., 1981. Altruism in the family and selfishness in the market. Economica 48, 1–15.

Bentham, J., 1879. Introduction to the Principles of Morals and Legislation. Oxford: Clarendon Press.

Berry, S., Levinsohn, J., Pakes, A., 1995. Automobile prices in market equilibrium. Econometrica 63, 841–890.

Bhattacharya, D., 2014. Nonparametric welfare analysis for discrete choice. Tech. rep., Oxford University.

Bilancini, E., 2011. On the rationalizability of observed consumers' choices when preferences depend on budget sets and (potentially) on anything else. Journal of Economics 102, 275–286.

Block, H. D., Marschak, J., 1959. Random orderings and stochastic theories of response. Discussion Paper 66, Cowles Foundation.

Blow, L., Browning, M., Crawford, I., 2008. Revealed preference analysis of characteristics models. Review of Economic Studies 75, 371–389.

Blundell, R., 2005. Presidential address: How revealing is revealed preference? Journal of the European Economic Association 3, 211–235.

Blundell, R., Browning, M., Cherchye, L., Crawford, I., De Rock, B., Vermeulen, F., 2012. Sharp for SARP: Nonparametric bounds on the behavioral and welfare effects of price changes. American Economic Journal: Microeconomics forthcoming (12.11).

Blundell, R., Browning, M., Crawford, I., 2003. Nonparametric engel curves and revealed preference. Econometrica 71, 205–240.

Blundell, R., Browning, M., Crawford, I., 2007. Improving revealed preference bounds on demand responses. International Economic Review 48, 1227–1244.

Blundell, R., Browning, M., Crawford, I., 2008. Best nonparametric bounds on demand responses. Econometrica 76, 1227–1262.

Blundell, R., Chiappori, P., Meghir, C., 2005. Collective labour supply with children. Journal of Political Economy 115, 1277–1306.

Blundell, R., Horowitz, J., Parey, M., 2013. Nonparametric estimation of a heterogeneous demand function under the Slutsky inequality restriction. Tech. rep., cemmap.

Blundell, R., Kristensen, D., Matzkin, R., 2014. Bounding quantile demand functions using revealed preference inequalities. Journal of Econometrics 179, 112–127.

Bolton, G. E., Ockenfels, A., 2000. Erc: A theory of equity, reciprocity, and competition. American Economic Review 90, 166–193.

Borkenau, P., Liebler, A., 1993. Convergence of stranger ratings of personality and intelligence with self-ratings, partner ratings, and measured intelligence. Journal of Personality and Social Psychology 65, 546–553.

Bronars, S. G., 1987. The power of nonparametric tests of preference maximization. Econometrica 55, 693–698.

Brown, B. W., Walker, M. B., 1989. The random utility hypothesis and inference in demand systems. Econometrica 57, 815–829.

Browning, M., 1989. A nonparametric test of the life-cycle rational expectations hypothesis. International Economic Review 30, 979–992.

Browning, M., Chiappori, P.-A., 1998. Efficient intra-household allocations: a general characterization and empirical tests. Econometrica 66, 1241–1278.

Bruyneel, S., Cherchye, L., Cosaert, S., De Rock, B., Dewitte, S., 2012a. Are the smart kids more rational? Tech. Rep. CES DP12.16, KU Leuven DPS 12.16.

Bruyneel, S., Cherchye, L., Cosaert, S., Rock, B. D., Dewitte, S., 2014. Measuring the willingness-to-pay for others' consumption: an application to joint decisions of children. Tech. rep., KU Leuven DPS 14.22.

Bruyneel, S., Cherchye, L., De Rock, B., 2012b. Collective consumption models with restricted bargaining weights: an empirical assessment based on experimental data. Review of Economics of the Household 10, 395–421.

Buhrmester, D., Goldfarb, J., Cantrell, D., 1992. Self-presentation when sharing with friends and nonfriends. Journal of Early Adolescence 12, 61–79.

Burghart, D. R., Glimcher, P. W., Lazzaro, S. C., 2012. An expected utility maximizer walks into a bar. Tech. rep., New York University.

Burks, S. V., Carpenter, J. P., Goette, L., Rustichini, A., 2009. Cognitive skills affect economic preferences, strategic behavior, and job attachment. Proceedings of the National Academy of Sciences 106, 7745–7750.

Caplin, A., Dean, M., 2014. Revealed preference, rational inattention and costly information acquisition. Tech. rep.

Caplin, A., Dean, M., Martin, D., 2011. Search and satisficing. American Economic Review 101, 2899–2922.

Carroll, J., 1993. Human cognitive abilities: A survey of factor-analytic studies. Cambridge University Press.

Charness, G., Rabin, M., 2002. Understanding social preferences with simple tests. Quarterly Journal of Economics 117, 817–869.

Cherchye, L., De Rock, B., Vermeulen, F., 2007. The collective model of household consumption: a nonparametric characterization. Econometrica 75, 553–574.

Cherchye, L., De Rock, B., Vermeulen, F., 2009. Opening the black box of intra-household decision-making: Theory and nonparametric empirical tests of general collective consumption models. Journal of Political Economy 117, 1074–1104.

Cherchye, L., De Rock, B., Vermeulen, F., 2011a. The revealed preference approach to collective consumption behavior: Testing and sharing rule recovery. Review of Economic Studies 78, 176–198.

Cherchye, L., De Rock, B., Vermeulen, F., 2012. Married with children: A collective labor supply model with detailed time use and intrahousehold expenditure information. American Economic Review 102, 3377–3405.

Cherchye, L., Demuynck, T., De Rock, B., 2011b. Revealed preference analysis of noncooperative household consumption. Economic Journal 121, 1073–1096.

Cherchye, L., Demuynck, T., De Rock, B., 2013a. Nash bargained consumption decisions: a revealed preference analysis. Economic Journal 123, 195–235.

Cherchye, L., Demuynck, T., De Rock, B., 2014. Revealed preference analysis for convex rationalization on nonlinear budget sets. Journal of Economic Theory 152, 224–236.

Cherchye, L., Demuynck, T., De Rock, B., De Witte, K., 2013b. Nonparametric analysis of multi-output production with joint inputs. Economic Journal forthcoming.

Chernozhukov, V., Hong, H., Tamer, E., 2007. Estimation and confidence regions for parameter sets in econometric models. Econometrica 75, 1243–1284.

Chernozhukov, V., Lee, S., Rosen, A. M., 2013. Intersection bounds: Estimation and inference. Econometrica 81, 667–737.

Chiappori, P., 1988. Rational household labor supply. Econometrica 56, 63–89.

Chiappori, P., 1992. Collective labor supply and welfare. Journal of Political Economy 100, 437–467.

Choi, S., Fisman, R., Gale, D., Kariv, S., 2007. Consistency and heterogeneity of individual behavior under uncertainty. American Economic Review 97, 1921–1938.

Choi, S., Kariv, S., Müller, W., Silverman, D., 2014. Who is (more) rational? American Economic Review 104, 1518–1550.

Christensen, L. R., Jorgenson, D. W., Lau, L. J., 1975. Transcendental logarithmic utility functions. American Economic Review 65, 367–383.

Cohen, M., Falmagne, J.-C., 1990. Random utility representation of binary choice probabilities: A new class of necessary conditions. Journal of Mathematical Psychology 34, 88–94.

Cohen, M. A., 1980. Random utility systems – the infinite case. Journal of Mathematical Psychology 22, 1–23.

Cosaert, S., 2013. Revealed preferences for diamond goods. Tech. rep., KU Leuven DPS 13.26.

Cosaert, S., Demuynck, T., 2013. Revealed preference theory for finite choice sets. Tech. rep., KU Leuven DPS 13.08.

Cosaert, S., Demuynck, T., 2014a. Nonparametric welfare and demand analysis with unobserved individual heterogeneity. Tech. rep., Maastricht University RM/14/010.

Cosaert, S., Demuynck, T., 2014b. Revealed preference theory for finite choice sets. Economic Theory forthcoming (DPS13.08).

Côté, S., Tremblay, R. E., Nagin, D., Zoccolillo, M., Vitaro, F., 2002. The development of impulsivity, fearfulness, and helpfulness during childhood: Patterns of consistency and change in the trajectories of boys and girls. Journal of Child Psychology and Psychiatry 43, 609–618.

Cox, J. C., 1997. On testing the utility hypothesis. Economic Journal 107, 1054–1078.

Cox, J. C., Friedman, D., Sadiraj, V., 2008. Revealed altruism. Econometrica 76, 31–69.

Crawford, I., 2010. Habits revealed. Review of Economic Studies 77, 1382–1402.

Damon, W., 1980. Patterns of change in children's social reasoning: A two-year longitudinal study. Child Development 51, 1010–1017.

Dauphin, A., El Lahga, A.-R., Fortin, B., Lacroix, G., 2011. Are children decision-makers within the household. Economic Journal 121, 871–903.

Dean, M., Martin, D., 2008. How consistent are your choice data? Tech. rep.

Deaton, A., Muellbauer, J., 1980. Economics and Consumer Behavior. Cambridge University Press.

Demuynck, T., 2014. Statistical inference for measures of predictive success. Working paper, 11.

Demuynck, T., Verriest, E., 2013. I'll never forget my first cigarette: A revealed preference analysis of the habits as durables model. International Economic Review 54, 717–738.

Diewert, W. E., 1973. Afriat and revealed preference theory. Review of Economic Studies 40, 419–425.

Dunbar, G. R., Lewbel, A., Pendakur, K., 2013. Children's resources in collective households: Identification, estimation, and an application to child poverty in Malawi. American Economic Review 103, 438–471.

Echenique, F., Lee, S., Shum, M., 2011. The money pump as a measure of revealed preference violations. Journal of Political Economy 119, 1201–1223.

Eisenberg, N., Fabes, R. A., Spinrad, T. L., 2007. Handbook of child psychology. Ch. Prosocial development.

Engelmann, D., Strobel, M., 2004. Inequality aversion, efficiency, and maximin preferences in simple distribution experiments. American Economic Review 94, 857–869.

Falmagne, J. C., 1978. A representation theorem for finite random scale systems. Journal of Mathematical Psychology 18, 52–72.

Farrell, L., Shields, M. A., 2007. Children as consumers: investigating child diary expenditure data. Canadian Journal of Economics 40, 445–467.

Fehr, E., Glätzle-Rützler, D., Sutter, M., 2013. The development of egalitarianism, altruism, spite and parochialism in childhood and adolescence. European Economic Review 64, 369–383.

Fehr, E., Schmidt, K. M., 1999. A theory of fairness, competition, and cooperation. Quarterly Journal of Economics 114, 817–868.

Février, P., Visser, M., 2004. A study of consumer behavior using laboratory data. Experimental Economics 7, 93–114.

Fishburn, P. C., 1978. Choice probabilities and choice functions. Journal of Mathematical Psychology 18, 205–219.

Fishburn, P. C., 1992. Induced binary probabilities and the linear ordering polytope: A status report. Mathematical Social Sciences 23, 67–80.

Fishburn, P. C., Falmagne, J.-C., 1989. Binary choice probabilities and rankings. Economics Letters 31, 113–117.

Fisman, R., Kariv, S., Markovits, D., 2007. Individual preferences for giving. American Economic Review 97, 1858–1876.

Forges, F., Iehlé, V., 2013. Essential data, budget sets and rationalization. Economic Theory 54, 449–461.

Forges, F., Iehlé, V., 2014. Afriat's theorem for indivisible goods. Journal of Mathematical Economics 54, 1–6.

Forges, F., Minelli, E., 2009. Afriat's theorem for general budget sets. Journal of Economic Theory 144, 135–145.

Fostel, A., Scarf, H. E., Todd, M. J., 2004. Two new proofs of Afriat's theorem. Economic Theory 24, 211–219.

Frank, R., Nagler, M., 2012. On the rationalizability of observed consumers' choices when preferences depend on budget sets: comment. Journal of Economics, 1–12.

Gardner, H., 1983. Frames of mind: the theory of multiple intelligences. Basic Books.

Giocoli, N., 2003. Modeling rational agents: From interwar economics to early modern game theory. Cheltenham: Edward Elgar.

Gorman, W., 1953. Community preference fields. Econometrica 21, 63–80.

Grüne, T., 2004. The problems of testing preference axioms with revealed preference theory. Analyse und Kritik 26, 382–397.

Guilbaud, G. T., 1953. Sur une difficulté de la théorie du risque. Colloques Internationaux dus Centre National de Recherche Scientifique, Econometrie 40, 19–25.

Gummerum, M., Keller, M., Takezawa, M., Mata, J., 2008. To give or not to give: children's and adolescents' sharing and moral negotiations in economic decision situations. Child Development 79, 562–576.

Hands, D. W., 2013. Foundations of contemporary revealed preference theory. Erkenntnis 78, 1081–1108.

Harbaugh, W. T., Krause, K., Berry, T. R., 2001. GARP for kids: On the development of rational choice behavior. American Economic Review 91, 1539–1545.

Harbaugh, W. T., Krause, K., Liday, S., 2003. Bargaining by children. Tech. rep., University of Oregon.

Harbaugh, W. T., Krause, K., Vesterlund, L., 2002. Risk attitudes of children and adults: choices over small and large probability gains and losses. Experimental Economics 5, 53–84.

Harbaugh, W. T., Krause, K., Vesterlund, L., 2007. Learning to bargain. Journal of Economic Psychology 28, 127–142.

Härdle, W., Jerison, M., 1988. The evolution of engel curves over time. Tech. rep., University of Bonn.

Härdle, W., Marron, J., 1990. Semiparametric comparison of regression curves. Annals of Statistics 18, 63–89.

Hausman, J. A., Newey, W. K., 2013. Individual heterogeneity and average welfare. Tech. rep., cemmap.

Heffetz, O., 2011. A test of conspicuous consumption: visibility and income elasticities. Review of Economics and Statistics 93, 1101–1117.

Heffetz, O., Shayo, M., 2009. How large are non-budget-constraint effects of prices on demand. American Economic Journal: Applied Economics 1, 170–199.

Hicks, J., 1946. Value and Capital. Oxford: Clarendon Press.

Hicks, J., Allen, R., 1934. A reconsideration of the theory of value, parts i and ii. Economica 2, 196–219.

Hoderlein, S., 2011. How many consumers are rational. Journal of Econometrics 164, 294–309.

Hoderlein, S., Stoye, J., 2013. Testing stochastic rationality and predicting stochastic demand: The case of two goods. Tech. rep.

Hoderlein, S., Stoye, J., 2014. Revealed preference in a heterogeneous population. Review of Economics and Statistics 96, 197–213.

Hoderlein, S., Vanhems, A., 2013. Estimating the distribution of welfare effects using quantiles. Tech. rep.

Hoge, R. D., Coladarci, T., 1989. Teacher-based judgments of academic achievement: A review of literature. Review of Educational Research 59, 297–313.

Houthakker, H. S., 1950. Revealed preference and the utility function. Economica 17, 159–174.

Houtman, M., Maks, J. A. H., 1985. Determining all maximal data subsets consistent with revealed preference. Kwantitatieve Methoden 19, 89–104.

Huck, S., Rasul, I., 2008. Testing consumer theory in the field: Private consumption versus charitable goods. Tech. rep., University College London.

Imbens, G. W., Manski, C. F., 2004. Confidence interval for partially identified parameters. Econometrica 72, 1845–1857.

Jevons, W. S., 1879. Theory of Political Economy. London: Macmillan.

John, D. R., 1999. Consumer socialization of children: a retrospective look at twenty-five years of research. Journal of Consumer Research 26, 183–213.

Kagan, S., Madsen, M., 1972. Rivalry in anglo-american and mexican children of two ages. Journal of Personality and Social Psychology 24, 214–220.

Kalman, P., 1968. Theory of consumer behavior when prices enter the utility function. Econometrica 36, 497–510.

Kawaguchi, K., 2012. Nonparametric test of axioms of revealed stochastic preference. Tech. rep., LSE.

Keita, L., 2012. Revealed preference theory, rationality, and neoclassical economics: Science or ideology. Africa Development 4, 73–116.

Kitamura, Y., Stoye, J., 2013. Nonparametric analysis of random utility models. CeMMap Working Paper CWP36/.13.

Leontief, W., 1936. Composite commodities and the problem of index numbers. Econometrica 4, 39–59.

Lewbel, A., 1996. Aggregation without separability: A generalized composite commodity theorem. American Economic Review 86, 524–543.

Lewbel, A., 2001. Demand systems with and without errors. American Economic Review 91, 611–618.

Li, Q., Racine, S., 2007. Nonparametric Econometrics. Princeton University Press, Princeton.

List, J. A., Millimet, D. L., 2008. The market: Catalyst for rationality and filter of irrationality. The B.E. Journal of Economic Analysis and Policy 8, 1–47.

Little, I. M. D., 1949. A reformulation of the theory of consumer's behaviour. Oxford Economic Papers 1, 90–99.

Lonnqvist, J. E., Vainikainen, M. P., Verkasalo, M., 2012. Teacher and parent ratings of seven-year-old children's personality and psychometrically assessed cognitive ability. European Journal of Personality 26, 504–514.

Lundberg, S., Romich, J., Tsang, K., 2009. Decision-making by children. Review of Economics of the Household 7, 1–30.

Mandel, B., 2009. Art as an investment and conspicuous consumption good. American Economic Review 99, 1653–1663.

Manski, C., 2001. Designing programs for heterogeneous populations: the value of covariate information. American Economic Review, Papers & Proceedings 91, 103–106.

Manski, C., 2007. Partial identification of counterfactual choice probabilities. International Economic Review 48, 1393–1410.

Marschak, J., 1960. Stanford Symposium on Mathematical Methods in the Social Sciences. Stanford University Press, Ch. Binary Choice Constraints on Random Utility Indicators.

Mattei, A., 2000. Full-scale real test of consumer behavior using experimental data. Journal of Economic Behavior and Organization 43, 487–497.

Matzkin, R. L., 1991. Axioms of revealed preference for nonlinear choice sets. Econometrica 59, 1779–1786.

McElroy, M. B., 1987. Additive general error models for production, cost, and derived demand or share systems. Journal of Political Economy 95, 737–757.

McFadden, D., 1975. Tchebycheff bounds for the space of agent characteristics. Journal of Mathematical Economics 2, 225–242.

McFadden, D., Richter, M. K., 1971. On the extension of a set function on a set of events to a probability on the generated boolean σ-algebra. Tech. rep., University of California, Berkeley.

McFadden, D. L., 2005. Revealed stochastic preference: a synthesis. Economic Theory 26, 245–264.

Murnighan, J. K., Saxon, M. S., 1998. Ultimatum bargaining by children and adults. Journal of Economic Psychology 19, 415–445.

Newey, W. K., Powell, J. L., Vella, F., 1999. Nonparametric estimation of triangular simultaneous equations models. Econometrica 67, 565–603.

Ng, Y., 1987. Diamonds are a government's best friend: Burden-free taxes on goods valued for their values. American Economic Review 77, 186–191.

Ng, Y., 1993. Mixed diamond goods and anomalies in consumer theory: Upward-sloping compensated demand curves with unchanged diamondness. Mathematical Social Sciences 25, 287–293.

Papke, L. E., Wooldridge, J. M., 1996. Econometric methods for fractional response variables with an application to 401(k) plan participation rates. Journal of Applied Econometrics 11, 619–632.

Pareto, V., 1927. Manual of Political Economy. New York: Augustus M. Kelly Publishers.

Piccione, M., Rubinstein, A., 2008. Luxury prices: An expository note. Japanese Economic Review 59, 127–132.

Pilgrim, C., Rueda-Riedle, A., 2002. The importance of social context in cross-cultural comparisons: First graders in columbia and the united states. Journal of Genetic Psychology 163, 283–295.

Pinske, C., Robinson, P., 1995. Pooling Nonparametric Estimates of Regression Functions with a Similar Shape. Oxford: Basil Blackwell, Ch. Advances in Econometrics and Quantitative Economics, pp. 172–195.

Polisson, M., Quah, J. K. H., 2013. Revealed preference in a discrete consumption space. American Economic Journal: Microeconomics 5, 28–34.

Politis, D. N., Romano, J. P., Wolf, M., 1999. Subsampling. Springer.

Pollak, R. A., 1971. The theory of the cost of living index. Research Discussion Paper 11, Bureau of Labor Statistics, Washington DC.

Pollak, R. A., 1977. Price dependent preferences. American Economic Review 67, 64–75.

Pollak, R. A., 1990. Houthakker's contributions to economics. Journal of Economic Perspectives 4, 141–156.

Rabin, M., 2002. A perspective on psychology and economics. European Economic Review 46, 657–685.

Rabin, M., 2013. An approach to incorporating psychology into economics. American Economic Review 103, 617–622.

Richter, M. K., 1966. Revealed preference theory. Econometrica 34, 635–645.

Rose, H., 1958. Consistency of preference: The two-commodity case. Review of Economic Studies 25, 124–125.

Samuelson, P. A., 1938. A note on the pure theory of consumer's behavior. Economica 5, 61–71.

Samuelson, P. A., 1948. Consumption theory in terms of revealed preference. Economica 15, 243–253.

Selten, R., 1991. Properties of a measure of predictive success. Mathematical Social Sciences 21, 153–167.

Sen, A. K., 1971. Choice functions and revealed preference. Review of Economic Studies 38, 307–317.

Sher, I., il Kim, K., Fox, J. T., Bajari, P., 2011. Partial identification of heterogeneity in preference orderings over discrete choices. Tech. rep., NBER.

Sippel, R., 1997. An experiment on the pure theory of consumer's behaviour. Economic Journal 107, 1431–1444.

Slutsky, E., 1915. Sulla teoria del bilancio del consonatore. Giornale degli Economisti 51, 1–26.

Sobel, J., 2005. Interdependent preferences and reciprocity. Journal of Economic Literature 43, 392–436.

Sternberg, R. J., 1985. Beyond IQ: a triarchic theory of human intelligence. Cambridge University Press.

Stigler, G. J., Becker, G. S., 1977. De gustibus non est disputandum. American Economic Review 67, 76–90.

Stoye, J., 2009. More on confidence intervals for partially identified parameters. Econometrica 77, 1299–1315.

Suzumura, K., 1976. Rational choice and revealed preference. Review of Economic Studies 43, 149–158.

Thaler, R., Sunstein, C., 2003. Libertarian paternalism. American Economic Review: Papers & Proceedings 93, 175–179.

Toda, M., Shinotsuka, H., McClintock, C. G., Stech, F. J., 1978. Development of competitive behavior as a function of culture, age, and social comparison. Journal of Personality and Social Psychology 36, 825–839.

Train, K., 2009. Discrete Choice Methods with Simulation. Cambridge University Press.

Varian, H. R., 1982. The nonparametric approach to demand analysis. Econometrica 50, 945–974.

Varian, H. R., 1983. Non-parametric tests of consumer behavior. Review of Economic Studies 50, 99–110.

Varian, H. R., 1990. Goodness-of-fit in optimizing models. Journal of Econometrics 46, 125–140.

Varian, H. R., 1993. Goodness-of-fit for revealed preference tests. Tech. rep., University of Michigan.

Varian, H. R., 2006. Samuelsonian economics and the 21st century. Oxford University Press, Ch. Revealed Preference.

Yatchew, A., Griliches, Z., 1985. Specification error in probit models. Review of Economics and Statistics 67, 134–139.

Yatchew, A. J., 1985. A note on nonparametric tests of consumer behavior. Economics Letters 18, 45–48.

List of doctoral dissertations

Doctoral dissertations from the Faculty of Business and Economics, see:

http://www.kuleuven.ac.be/doctoraatsverdediging/archief.htm